PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :		(11) International Publication Number:
C12P 21/06, C12N 15/00	A1	WO 91/10741
		(43) International Publication Date: 25 July 1991 (25.07.91)

(21) International Application Number: PCT/US91/00245

(22) International Filing Date: 11 January 1991 (11.01.91)

(30) Priority data:

466,008	12 January 1990 (12.01.90)	US
610,515	8 November 1990 (08.11.90)	US

(71) Applicant: CELL GENESYS, INC. [US/US]; 344A Lakeside Drive, Foster City, CA 94404 (US).

(72) Inventors: KUCHERLAPATI, Raju ; 8 Gracie Lane, Darien, CT 06820 (US). JAKOBOVITS, Aya ; 2021 Monterey Avenue, Menlo Park, CA 94025 (US).

(74) Agents: ROWLAND, Bertram, I. et al.; Cooley Godward Castro Huddleson & Tatum, Five Palo Alto Square, 4th Floor, Palo Alto, CA 94306 (US).

(81) Designated States: AT (European patent), AU, BE (European patent), CA, CH (European patent), DE (European patent), DK (European patent), ES (European patent), FR (European patent), GB (European patent), GR (European patent), IT (European patent), JP, KR, LU (European patent), NL (European patent), NO, SE (European patent).

Published
With international search report.

(54) Title: GENERATION OF XENOGENEIC ANTIBODIES

(57) Abstract

The subject invention provides non-human mammalian hosts characterized by inactivated endogenous Ig loci and functional human Ig loci for response to an immunogen to produce human antibodies or analogs thereof. The hosts are produced by repetitive transformations of embryonic stem cells by homologous recombination, preferably in conjunction with breeding. Different strategies are employed for recombination of the human loci randomly or at analogous host loci.

Attorney Docket No: 10271-054-999
Application Serial No.: 10/091,268
Reference B37

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	ES	Spain	MG	Madagascar
AU	Australia	FI	Finland	ML	Mali
BB	Barbados	FR	France	MN	Mongolia
BE	Belgium	GA	Gabon	MR	Mauritania
BF	Burkina Faso	GB	United Kingdom	MW	Malawi
BG	Bulgaria	GN	Guinea	NL	Netherlands
BJ	Benin	GR	Greece	NO	Norway
BR	Brazil	HU	Hungary	PL	Poland
CA	Canada	IT	Italy	RO	Romania
CF	Central African Republic	JP	Japan	SD	Sudan
CG	Congo	KP	Democratic People's Republic of Korea	SE	Sweden
CH	Switzerland			SN	Senegal
CI	Côte d'Ivoire	KR	Republic of Korea	SU	Soviet Union
CM	Cameroon	LI	Liechtenstein	TD	Chad
CS	Czechoslovakia	LK	Sri Lanka	TG	Togo
DE	Germany	LU	Luxembourg	US	United States of America
DK	Denmark	MC	Monaco		

GENERATION OF XENOGENEIC ANTIBODIES

CROSS-REFERENCE TO RELATED APPLICATIONS

This application is a continuation-in-part of Application Serial No. 07/466,008 filed January 12, 1990.

INTRODUCTION

Technical Field

The field of this invention is the production of xenogeneic specific binding proteins in a viable mammalian host.

Background

Monoclonal antibodies find use in both diagnosis and therapy. Because of their ability to bind to a specific epitope, they can be uniquely used to identify molecules carrying that epitope or may be directed, by themselves or in conjunction with another moiety, to a specific site for diagnosis or therapy.

Monoclonal antibodies comprise heavy and light chains which join together to define a binding region for the epitope. Each of the chains is comprised of a variable region and a constant region. The constant region amino acid sequence is specific for a particular isotype of the antibody, as well as the host which produces the antibody.

Because of the relationship between the sequence of the constant region and the species from which the antibody is produced, the introduction of a xenogeneic antibody into the vascular system of the host can produce an immune response. Where the xenogeneic antibody may be introduced repetitively, in the case of chronic diseases, it becomes impractical to administer the antibody, since it will be rapidly

destroyed and may have an adverse effect. There have been, therefore, many efforts to provide a source of syngeneic or allogeneic antibodies. One technique has involved the use of recombinant DNA technology where the genes for the heavy and light chains from the host were identified and the regions encoding the constant region isolated. These regions were then joined to the variable region encoding portion of other immunoglobulin genes from another species directed to a specific epitope.

While the resulting chimeric partly xenogeneic antibody is substantially more useful than using a fully xenogeneic antibody, it still has a number of disadvantages. The identification, isolation and joining of the variable and constant regions requires substantial work. In addition, the joining of a constant region from one species to a variable region from another species may change the specificity and affinity of the variable regions, so as to lose the desired properties of the variable region. Also, there are framework and hypervariable sequences specific for a species in the variable region. These framework and hypervariable sequences may result in undesirable antigenic responses.

It would therefore be more desirable to produce allogeneic antibodies for administration to a host by immunizing the host with an immunogen of interest. For primates, particularly humans, this approach is not practical. The human antibodies which have been produced have been based on the adventitious presence of an available spleen, from a host which had been previously immunized to the epitope of interest. While human peripheral blood lymphocytes may be employed for the production of monoclonal antibodies, these have not been particularly successful in fusions and have usually led only to IgM. Moreover, it is particularly difficult to generate a human antibody response against a human protein, a desired target in

many therapeutic and diagnostic applications. There is, therefore, substantial interest in finding alternative routes to the production of allogeneic antibodies for humans.

Relevant Literature

Thomas and Capecchi, Cell, 51, 503-512, 1987. Koller and Smithies, Proc. Natl. Acad. Sci. USA, 86, 8932-8935, 1989, describe inactivating the B2microglobulin locus by homologous recombination in embryonic stem cells. Berman et al., EMBO J. 7, 727-738, 1988, describe the human Ig VH locus. Burke, et al., Science, 236, 806-812, 1987, describe yeast artificial chromosome vectors. See also, Garza et al., Science, 246, 641-646, 1989, and Brownstein et al., Science, 244, 1348-1351, 1989. Sakano, et al., describe a diversity segment of the immunoglobulin heavy chain genes. Sakano et al., Nature, 290, 562-565, 1981. Tucker et al., Proc. Natl. Acad. Sci. USA, 78, 7684-7688, 1981, describe the mouse IgA heavy chain gene sequence. Blankenstein and Kruwinkel Eur. J. Immunol., 17, 1351-1357, 1987, describe the mouse variable heavy chain region. See also, Joyner et al., Nature, 338, 153-155, 1989, Traver et al., Proc. Nat. Acad. Sci. USA 86, 5898-5902, 1989, and Panchis et al., Proc. Nat. Acad. Sci. USA , 87, 5109-5113, 1990.

SUMMARY OF THE INVENTION

Xenogeneic specific binding proteins are produced in a non-primate viable mammalian host by immunization of the mammalian host with an appropriate immunogen.

The host is characterized by: (1) being incapable of producing endogenous immunoglobulin; (2) an exogenous immunoglobulin locus comprising at least one immunoglobulin constant region, or protein thereof, immunoglobulin sequences providing for the components of the variable region of at least one of

the light and heavy chains, and at least one intron with appropriate splicing sites for excision and assembly of a functional immunoglobulin subunit. Thus, the mammalian host will comprise at least one xenogeneic constant region or protein thereof capable of being spliced to a functional J region of an endogenous or exogenous immunoglobulin locus, may have an entire immunoglobulin locus of the host substituted by a portion or an entire xenogeneic immunoglobulin locus, or may have a xenogeneic immunoglobulin locus inserted into a chromosome of the host cell and an inactivated endogenous immunoglobulin region. These various alternatives will be achieved, at least in part, by employing homologous recombination at the immunoglobulin loci for the heavy and light chains.

DESCRIPTION OF SPECIFIC EMBODIMENTS

Novel transgenic mammalian hosts, other than primates, particularly other than human, are provided, where the host is capable of mounting an immune response to an immunogen, where the response produces antibodies having primate, particularly human, constant and/or variable regions or such other effector peptide sequences of interest. The hosts are characterized by being capable of producing xenogeneic or modified antibodies as a result of substitution and/or inactivation of the endogenous immunoglobulin subunit encoding loci. The modifications retain at least a portion of the constant regions which provide for assembly of the variable region binding site bonded at the C-terminus to a functional peptide. The functional peptide may take many forms or conformations and may serve as an enzyme, growth factor, binding protein, ligand, cytokine, effector protein, chelating proteins, etc. The antibodies may be any isotype, e.g., IgA, D, E, G or M or subtypes within the isotype.

A number of strategies may be employed to achieve the desired transgenic hosts. Various transgenic hosts may be employed, particularly murine, lagomorpha, ovine, porcine, equine, canine, feline, or the like, normally other than primate. For the most part, mice have been used for the production of B-lymphocytes for immortalization for the production of antibodies. Since mice are easy to handle, can be produced in large quantities, and are known to have an extensive immune repertoire, mice will usually be the animal of choice. Therefore, in the following discussion, the discussion will refer to mice, but it should be understood that other animals, particularly mammals, may be readily substituted for the mice, following the same procedures.

In one strategy, as individual steps, the human heavy and light chain immunoglobulin gene complexes are introduced into the mouse germ line and in a separate step the corresponding mouse genes are rendered non-functional. Human heavy and light chain genes are reconstructed in an appropriate eukaryotic or prokaryotic microorganism and the resulting DNA fragments can be introduced into pronuclei of fertilized mouse oocytes or embryonic stem cells. Inactivation of the endogenous mouse immunoglobulin loci is achieved by targeted disruption of the appropriate loci by homologous recombination in mouse embryonic stem cells. In each case chimeric animals are generated which are derived in part from the modified embryonic stem cells and are capable of transmitting the genetic modifications through the germ line. The mating of mouse strains with human immunoglobulin loci to strains with inactivated mouse loci will yield animals whose antibody production is purely human.

In the next strategy, fragments of the human heavy and light chain immunoglobulin loci are used to directly replace the corresponding mouse loci by

homologous recombination in mouse embryonic stem cells. This is followed by the generation of chimeric transgenic animals in which the embryonic stem cell-derived cells contribute to the germ line.

These strategies are based on the known organization of the immunoglobulin chain loci in a number of animals, since the organization, relative location of exons encoding individual domains, and location of splice sites and transcriptional elements, is understood to varying degrees. In the human, the immunoglobulin heavy chain locus is located on chromosome 14. In the 5' - 3' direction of transcription, the locus comprises a large cluster of variable region genes (V_H), the diversity (D) region genes, followed by the joining (J_H) region genes and the constant (C_H) gene cluster. The size of the locus is estimated to be about 2,500 kilobases (kb). During B-cell development, discontinuous gene segments from the germ line IgH locus are juxtaposed by means of a physical rearrangement of the DNA. In order for a functional heavy chain Ig polypeptide to be produced, three discontinuous DNA segments, from the V_H, D, and J_H regions must be joined in a specific sequential fashion; V_H to DJ_H, generating the functional unit V_HDJ_H. Once a V_HDJ_H has been formed, specific heavy chains are produced following transcription of the Ig locus, utilizing as a template the specific $V_HDJ_HC_H$ unit comprising exons and introns. There are two loci for Ig light chains, the κ locus on human chromosome 2 and the λ locus on human chromosome 22. The structure of the IgL loci is similar to that of the IgH locus, except that the D region is not present. Following IgH rearrangement, rearrangement of a light chain locus is similarly accomplished by V_L and J_L joining of the κ or λ chain. The sizes of the λ and κ loci are each approximately 1000 kb. Expression of rearranged IgH and an Igκ or Igλ light chain in a

particular B-cell allows for the generation of antibody molecules.

In order to isolate, clone and transfer the IgH_{hu} locus, a yeast artificial chromosome may be employed. The entire IgH_{hu} locus can be contained within one or a few yeast artificial chromosome (YAC) clones. The same is true for the Ig light chain loci. Subsequent introduction of the appropriate heavy chain or light chain YAC clones into recipient yeast allows for the reconstitution of intact germ line Ig loci by homologous recombination between overlapping regions of homology. In this manner, the isolation of DNA fragments encoding the human Ig chain is obtained.

In order to obtain a broad spectrum of high affinity antibodies, it is not necessary that one include the entire V region. Various V region gene families are interspersed within the V region cluster. Thus, by obtaining a subset of the known V region genes of the human heavy and light chain Ig loci (Berman et al., EMBO J. (1988) 7: 727-738) rather than the entire complement of V regions, the transgenic host may be immunized and be capable of mounting a strong immune response and provide high affinity antibodies. In this manner, relatively small DNA fragments of the chromosome may be employed, for example, a reported 670 kb fragment of the Ig_{Hu} locus is shown in Figure 1b. This NotI-NotI restriction fragment would serve to provide a variety of V regions, which will provide increased diversity by recombining with the various D and J regions and undergoing somatic mutation.

In order to provide for the production of human antibodies in a xenogeneic host, it is necessary that the host be competent to provide the necessary enzymes and other factors involved with the production of antibodies, while lacking competent endogenous genes for the expression of heavy and light subunits of immunoglobulins. Thus, those enzymes and other

factors associated with germ line rearrangement, splicing, somatic mutation, and the like, will be functional in the xenogeneic host. What will be lacking is a functional natural region comprising the various exons associated with the production of endogenous immunoglobulin subunits.

The human DNA may be introduced into the pronuclei of fertilized oocytes or embryonic stem cells. The integration may be random or homologous depending on the particular strategy to be employed. Thus, by using transformation, using repetitive steps or in combination with breeding, transgenic animals may be obtained which are able to produce human antibodies in the substantial absence of light or heavy host immunoglobulin subunits.

To inactivate the host immunoglobulin loci, homologous recombination may be employed, where DNA is introduced at the immunoglobulin heavy chain and light chain loci which inhibits the production of endogenous immunoglobulin subunits. Since there are two heavy chain alleles and two light chain loci, each with two alleles, although one may choose to ignore the λ loci, there will have to be multiple transformations which result in inactivation of each of the alleles. (By transformation is intended any technique for introducing DNA into a viable cell, such as conjugation, transformation, transfection, transduction, electroporation, lipofection, biolistics, or the like.) Homologous recombination may be employed to functionally inactivate each of the loci, by introduction of the homologous DNA into embryonic stem cells, followed by introduction of the modified cells into recipient blastocysts. Subsequent breeding allows for germ line transmission of the inactivated locus. One can therefore choose to breed heterozygous offspring and select for homozygous offspring from the heterozygous parents or again one may use the embryonic stem cell for homologous

recombination and inactivation of the comparable locus.

The number of transformation steps may be reduced by providing at least a fragment of the human immunoglobulin subunit locus for homologous recombination with the analogous endogenous immunoglobulin, so that the human locus is substituted for at least a part of the host immunoglobulin locus, with resulting inactivation of the host immunoglobulin subunit locus. Of particular interest is the use of transformation for a single inactivation, followed by breeding of the heterozygous offspring to produce a homozygous offspring. Where the human locus is employed for substitution or insertion into the host locus for inactivation, the number of transformations may be limited to three transformations and as already indicated, one may choose to ignore the less used λ locus and limit the transformations to two transformations. Alternatively, one may choose to provide for inactivation as a separate step for each locus, employing embryonic stem cells from offspring which have previously had one or more loci inactivated. In the event only transformation is used and the human locus is integrated into the host genome in random fashion, a total of eight transformations may be required.

For inactivation, any lesion in the target locus resulting in the prevention of expression of an immunoglobulin subunit of that locus may be employed. Thus, the lesion may be in a region comprising the enhancer, e.g., 5' upstream or intron, in the V, J or C regions, and with the heavy chain, the opportunity exists in the D region, or combinations thereof. Thus, the important factor is that Ig germ line gene rearrangement is inhibited, or a functional message encoding the immunoglobulin subunit cannot be produced, either due to failure of transcription, failure of processing of the message, or the like.

Preferably, when one is only interested in inactivating the immunoglobulin subunit locus, the lesion will be introduced into the J region of the immunoglobulin subunit locus. Thus, one produces a construct which lacks a functional J region and may comprise the sequences of the J region adjacent to and upstream and/or downstream from the J region or comprises all or part of the region with an inactivating insertion in the J region. The insertion may be 50 bp or more, where such insertion results in disruption of formation of a functional mRNA. Desirably, the J region in whole or substantial part, usually at least about 75% of the locus, preferably at least about 90% of the locus, is deleted. If desired, the lesion between the two flanking sequences defining the homologous region may extend beyond the J region, into the variable region and/or into the constant region.

Desirably, a marker gene is used to replace the J region. Various markers may be employed, particularly those which allow for positive selection. Of particular interest is the use of G418 resistance, resulting from expression of the gene for neomycin phosphotransferase.

Upstream and/or downstream from the target gene construct may be a gene which provides for identification of whether a double crossover has occurred. For this purpose, the Herpes simplex virus thymidine kinase gene may be employed, since cells expressing the thymidine kinase gene may be killed by the use of nucleoside analogs such as acyclovir or gancyclovir, by their cytotoxic effects on cells that contain a functional HSV-tk gene. The absence of sensitivity to these nucleoside analogs indicates the absence of the HSV-thymidine kinase gene and, therefore, where homologous recombination has occurred, that a double crossover has also occurred.

While the presence of the marker gene in the genome will indicate that integration has occurred, it will still be necessary to determine whether homologous integration has occurred. This can be achieved in a number of ways. For the most part, DNA analysis will be employed to establish the location of the integration. By employing probes for the insert and then sequencing the 5' and 3' regions flanking the insert for the presence of the target locus extending beyond the flanking region of the construct or identifying the presence of a deletion, when such deletion has been introduced, the desired integration may be established.

The polymerase chain reaction (PCR) may be used with advantage in detecting the presence of homologous recombination. Probes may be used which are complementary to a sequence within the construct and complementary to a sequence outside the construct and at the target locus. In this way, one can only obtain DNA chains having both the primers present in the complementary chains if homologous recombination has occurred. By demonstrating the presence of the probes for the expected size sequence, the occurrence of homologous recombination is supported.

The construct may further include a replication system which is functional in the mammalian host cell. For the most part, these replication systems will involve viral replication systems, such as Simian virus 40, Epstein-Barr virus, polyoma virus, papilloma virus, and the like. Various transcriptional initiation systems may be employed, either from viruses or from mammalian genes, such as SV40, metallathionein-I and II genes, β-actin gene, adenovirus early and late genes, phosphoglycerate kinase gene, RNA polymerase II gene, or the like. In addition to promoters, wild-type enhancers may be employed to further enhance the expression of the marker gene.

In constructing the subject constructs for homologous recombination, a replication system for procaryotes, particularly E. coli, may be included, for preparing the construct, cloning after each manipulation, analysis, such as restriction mapping or sequencing, expansion and isolation of the desired sequence. Where the construct is large, generally exceeding about 50 kbp, usually exceeding 100 kbp, and usually not more than about 1000kbp, a yeast artificial chromosome (YAC) may be used for cloning of the construct.

Once a construct has been prepared and any undesirable sequences removed, e.g., procaryotic sequences, the construct may now be introduced into the target cell. Any convenient technique for introducing the DNA into the target cells may be employed. Techniques include spheroplast fusion, lipofection, electroporation, calcium phosphate-mediated DNA transfer or direct microinjection. After transformation or transfection of the target cells, target cells may be selected by means of positive and/or negative markers, as previously indicated, neomycin resistance and acyclovir or gancyclovir resistance. Those cells which show the desired phenotype may then be further analyzed by restriction analysis, electrophoresis, Southern analysis, PCR, or the like. By identifying fragments which show the presence of the lesion(s) at the target locus, one can identify cells in which homologous recombination has occurred to inactivate a copy of the target locus.

The above described process may be performed first with a heavy chain locus in an embryonic stem cell and then maturation of the cells to provide a mature fertile host. Then by breeding of the heterozygous hosts, a homozygous host may be obtained or embryonic stem cells may be isolated and transformed to inactivate the second IgH locus, and the process repeated until all the desired loci have

been inactivated. Alternatively, the light chain locus may be the first. At any stage, the human loci may be introduced.

As already indicated, the target locus may be substituted with the analogous human locus. In this way, the human locus will be placed substantially in the same region as the analogous host locus, so that any regulation associated with the position of the locus will be substantially the same for the human immunoglobulin locus. For example, by isolating the entire V_H gene locus (including V, D, and J sequences), or portion thereof, and flanking the human locus with sequences from the mouse locus, preferably sequences separated by at least about 5 kbp, in the host locus, preferably at least about 10 kbp in the host locus, one may insert the human fragment into this region in a recombinational event(s), substituting the human immunoglobulin locus for the variable region of the host immunoglobulin locus. In this manner, one may disrupt the ability of the host to produce an endogenous immunoglobulin subunit, while allowing for the promoter of the human immunoglobulin locus to be activated by the host enhancer and regulated by the regulatory system of the host.

Once the human loci have been introduced into the host genome, either by homologous recombination or random integration, and host animals have been produced with the endogenous immunoglobulin loci inactivated by appropriate breeding of the various transgenic or mutated animals, one can produce a host which lacks the native capability to produce endogenous immunoglobulin subunits, but has the capacity to produce human immunoglobulins with at least a significant portion of the human repertoire.

The functional inactivation of the two copies of each of the three host Ig loci, where the host contains the human IgH and the human Ig κ and/or λ loci would allow for the production of purely

human antibody molecules without the production of host or host/human chimeric antibodies. Such a host strain, by immunization with specific antigens, would respond by the production of mouse B-cells producing specific human antibodies, which B-cells could be fused with mouse myeloma cells or be immortalized in any other manner for the continuous stable production of human monoclonal antibodies.

The subject methodology and strategies need not be limited to producing complete immunoglobulins, but provides the opportunity to provide for regions joined to a portion of the constant region, e.g., C_{H1}, C_{H2}, C_{H3}, or C_{H4}, or combination thereof. Alternatively, one or more of the exons of the C_H and C_κ or C_λ regions may be replaced or joined to a sequence encoding a different protein, such as an enzyme, e.g., plasminogen activator, superoxide dismutase, etc.; toxin A chain, e.g., ricin, abrin, diphtheria toxin, etc.; growth factors; cytotoxic agent, e.g., TNF, or the like. See, for example, WO 89/07142; WO 89/09344; and WO 88/03559. By inserting the protein of interest into a constant region exon and providing for splicing of the variable region to the modified constant region exon, the resulting binding protein may have a different C-terminal region from the immunoglobulin. By providing for a stop sequence with the inserted gene, the protein product will have the inserted protein as the C-terminal region. If desired, the constant region may be entirely substituted by the other protein, by providing for a construct with the appropriate splice sites for joining the variable region to the other protein.

The antibodies or antibody analog producing B-cells from the transgenic host may be used for fusion to a mouse myeloid cell to produce hybridomas or immortalized by other conventional process, e.g., transfection with oncogenes. These immortalized cells

may then be grown in continuous culture or introduced into the peritoneum of a compatible host for production of ascites.

The subject invention provides for the production of polyclonal human anti-serum or human monoclonal antibodies or antibody analogs. Where the mammalian host has been immunized with an immunogen, the resulting human antibodies may be isolated from other proteins by using an affinity column, having an Fc binding moiety, such as protein A, or the like.

For producing animals from embryonic stem cells, after transformation, the cells may be plated onto a feeder layer in an appropriate medium, e.g. fetal bovine serum enhanced DMEM. Cells containing the construct may be detected by employing a selective medium and after sufficient time for colonies to grow, colonies may be picked and analyzed for the occurrence of integration or homologous recombination. As described previously, the PCR may be used, with primers within or without the construct sequence, but at the target locus.

Those colonies which show homologous recombination may then be used for embryo manipulation and blastocyst injection. Blastocysts may be obtained from females by flushing the uterus 3-5 days after ovulation. The embryonic stem cells may then be trypsinized and the modified cells added to a droplet containing the blastocyst. At least one and up to thirty modified embryonic stem cells may be injected into the blastocoel of the blastocyst. After injection, at least one and no more then about fifteen of the blastocysts are returned to each uterine horn of pseudo-pregnant females. Females are then allowed to go to term and the resulting litter screened for mutant cells having the construct.

The mammals may be any non-human, particularly non-primate mammal, such as laboratory animals, particularly small laboratory animals, such

as mice, rats, guinea pigs, etc., domestic animals, pets, etc.

The following examples are offered by way of illustration and not by way of limitation.

EXPERIMENTAL

Inactivation of the mouse heavy chain J genes

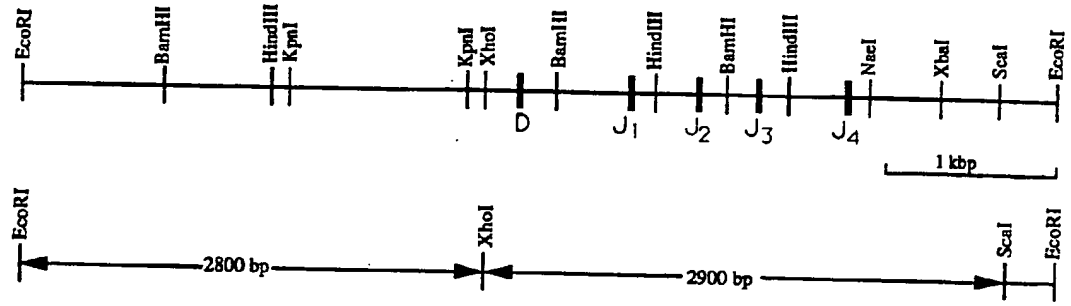
Construction of the inactivation vector

A 6.4 Kb EcoRI fragment, containing the mouse heavy chain J genes and flanking sequences, is cloned from a Balb/c mouse embryo genomic library using the probes described in Sakano et al., Nature 290:562-565, 1981. This fragment (mDJ) is inserted into EcoRI-digested pUC19 plasmid (pmDJ). A 2.9 Kb fragment, containing the 4 J genes, is deleted by XhoI-ScaI digestion (pmDδJNeo, see Chart 1). An 1150 bp XhoI-BamHI fragment, containing a neomycin-resistance gene driven by the Herpes simplex virus thymidine kinase gene (HSV-tk) promoter and a polyoma enhancer is isolated from pMC1Neo (Thomas and Capecchi, Cell, 51, 503-512, 1987). A synthetic adaptor is added onto this fragment to convert the BamHI end into a ScaI end and the resulting fragment is joined to the XhoI-ScaI pmDδJ to form the inactivation vector (pmDδJ.Neo) in which the 5′ to 3′ orientation of the neomycin and the heavy chain promoters is identical. This plasmid is linearized by NdeI digestion before transfection to ES cells. The sequences driving the homologous recombination event are 3 kb and 0.5 kb fragments, located 5′ and 3′ to the neomycin gene, respectively.

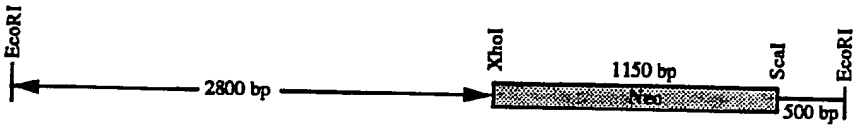
Chart 1

Mouse Heavy Chain J Genes Inactivation Vector

(A) Targeted mouse heavy chain J genes



(B) Inactivation vector mDΔJ.Neo



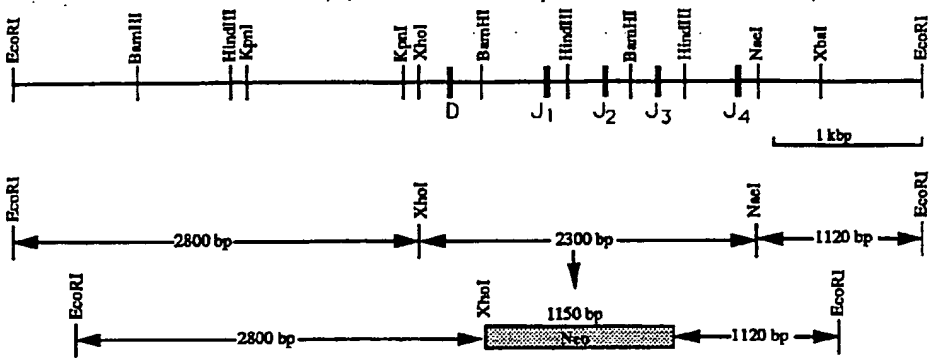
The ES cell line E14TG2a (Hooper et al., Nature, 326:292-295, 1987) is cultured on mitomycin-treated primary embryonic fibroblast-feeder layers essentially as described (Doetschman et al., J. Embryol. Exp. Morphol. 87:27-45, 1985). The embryonic fibroblasts are prepared from embryos from C57BL/6 females that are mated 14 to 17 days earlier with a male homozygous for a neomycin transgene (Gossler et al., PNAS 83:9065-9069, 1986). These cells are capable of growth in media containing G418. Electroporation conditions are described by (Boggs et al., Ex. Hematol. (NY) 149:988-994, 1986). ES cells are trypsinized, resuspended in culture media at a concentration of $4x10^7$/ml and electroporated in the presence of the targeting DNA at a concentration of 12nM in the first experiment and 5nM DNA in the second. A voltage of 300 V with a capacitance of 150-250 μF is found optimal with an electroporation cell of 5 mm length and 100 mm^2 cross-section. $5x10^6$ electroporated cells are plated onto mitomycin-treated fibroblasts in 100 mm dishes in the presence of Dulbecco's modified Eagle's media (DMEM) supplemented with 15% fetal bovine serum (FBS) and 0.1 mM 2-mercaptoethanol. The media is replaced 24 hr after electroporation with media containing 200 μg/ml G418.

ES colonies resulting 10-14 days after electroporation are picked with drawn out capillary pipettes for analysis using PCR. Half of each picked colony is saved in 24-well plates already seeded with mitomycin-treated feeder cells. The other halves, combined in pools of 3-4, are transferred to Eppendorf tubes containing approximately 0.5 ml of PBS and analyzed for homologous recombination by PCR. Conditions for PCR reactions are essentially as described (Kim and Smithies, Nucleic Acids Res. 16:8887-8893, 1988). After pelleting, the ES cells are resuspended in 5 μl of PBS and are lysed by the

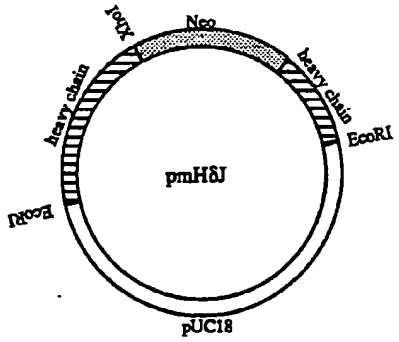
addition of 55 μl of H_2O to each tube. DNAses are inactivated by heating each tube at 95°C for 10 min. After treatment with proteinase K at 55°C for 30 min, 30 μl of each lysate is transferred to a tube containing 20 μl of a reaction mixture including PCR buffer: 1.5 μg of each primer, 3U of Taq polymerase, 10% DMSO, and dNTPs, each at 0.2 mM. The PCR expansion employs 55 cycles using a thermocycler with 65 seconds melt at 92°C and a 10 min annealing and extension time at 65°C. The two priming oligonucleotides are TGGCGGACCGCTATCCCCCAGGAC and TAGCCTGGGTCCCTCCTTAC, which correspond respectively to a region 650 bases 3' of the start codon of the neomycin gene and sequences located in the mouse heavy chain gene, 1100 bases 3' of the insertion site. 20 μl of the reaction mix is electrophoresed on agarose gels and transferred to nylon membranes (Zeta Bind). Filters are probed with a ^{32}P-labelled fragment of the 991 bp XbaI fragment of the J-C region.

Chart 2

(A) Targeted mouse heavy chain J genes

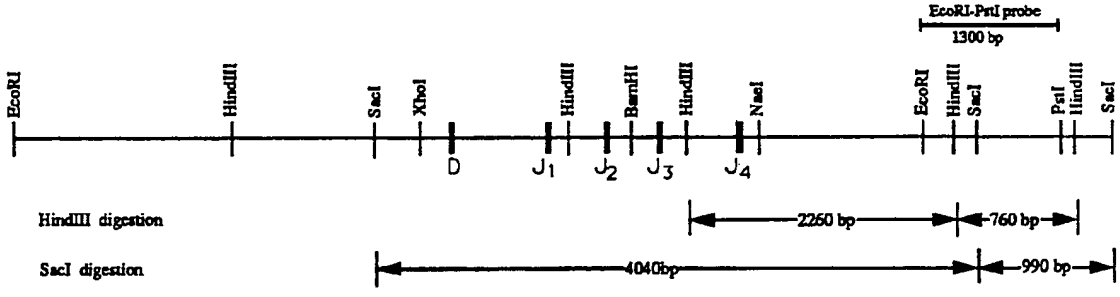


(B) Inactivation vector pmHδJ

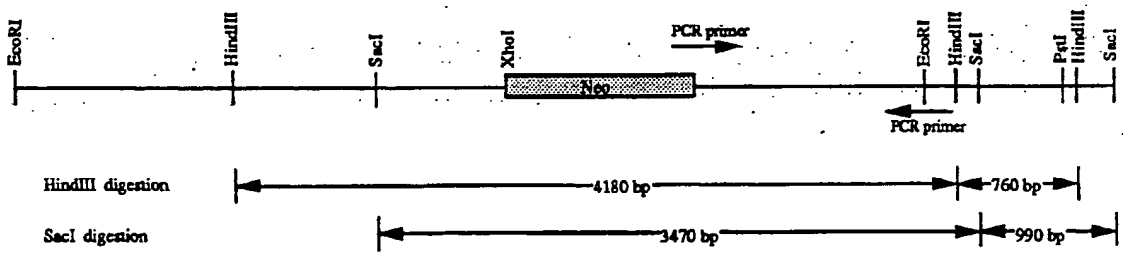


(C) Southern analysis of pmHδJ-targeted ES colonies

Wild type ES cell genome



Targeted ES cell genome



Inactivation of the mouse Ig heavy chain J genes in ES cells.

Construction of the inactivation vector

A 6.1-Kb EcoRI fragment, containing the mouse immunoglobulin heavy chain J region genes and flanking sequences, cloned from a Balb/c mouse embryo genomic library and inserted into pUC18 (pJH), was digested with XhoI and NaeI to delete an about 2.3 kbp fragment containing the four J genes (see Chart 2A). An about 1.1 kbp XhoI-BamHI fragment, blunted at the BamHI site, containing a neomycin resistance gene driven by the Herpes simplex virus thymidine kinase gene (HSV-tk) promoter and polyoma enhancer was isolated from pMC1Neo (Thomas and Capecchi, Cell, 51, 503-512, 1987). This fragment was inserted into the XhoI-NaeI deleted pJH to form the inactivation vector (pmHδJ, see Chart 2B), in which the transcriptional orientation of the neomycin and the heavy chain genes is the same. This plasmid was linearized by NdeI digestion before transfection to ES cells. The sequences driving the homologous recombination event are about 2.8 kbp and about 1.1 kbp fragments, located 5' and 3' to the neomycin gene, respectively.

Culturing, Electroporation, and Selection of ES cells

The ES cell line E14TG2a (Koller and Smithies, 1989, PNAS, USA, 86, 8932-8935) was cultured on mitomycin C-treated embryonic fibroblast feeder layers as described (Koller and Smithies, 1989, PNAS USA, 86, 8932-8935). ES cells were trypsinized, resuspended in HBS buffer (pH 7.05; 137mM NaCl, 5mM KCl, 2mM $CaCl_2$, 0.7mM Na_2HPO_4, 21mM HEPES pH7.1) at a concentration of $2x10^7$/ml and electroporated in the presence of 50 μg/ml of the linearized inactivation vector. Electroporation was carried out with a BioRad Gene Pulser using 240 volts and 500μF capacitance. $5x10^6$ electroporated cells were plated onto mitomycin C-treated fibroblasts in 100mm dishes in the presence

of Dulbecco's modified Eagle's media (DMEM) supplemented with 15% fetal bovine serum and 0.1 mM 2-mercaptoethanol. The media was replaced 24 hr after electroporation with media containing 200μg/ml G418. G418-resistant ES colonies resulting 12-14 days after electroporation were picked with drawn out capillary pipettes for analysis using the polymerase chain reaction (PCR). Half of each picked colony was transferred to an individual well of a 24-well plate, already seeded with mitomycin C-treated feeder cells. The other halves, combined in pools of four, were transferred to Eppendorf tubes containing 0.3 ml of PBS and cell lysates were prepared for PCR analysis as described by Joyner et al (Nature, 338:153-155, 1989). The PCR reaction included 5-20μl of the cell lyste, 1μM of each primer, 1.5u of Taq polymerase and 200μM of dNTPs. The PCR amplification employed 45 cycles using a thermal cycler (Perkin-Elmer Cetus), with 1 min. melt at 94°C, 2 min. annealing at 55°C, and 3 min. extension at 72°C. The two priming oligonucleotides are ACGGTATCGCCGCTCCCGAT and AGTCACTGTAAAGACTTCGGGTA, which correspond respectively to about 120 bases 5' of the BamHI site of the neomycin gene, and to the sequences located in the mouse heavy chain gene, about 160 bases 3' of the insertion site. Successful homologous recombination gives rise to an about 1.4 kbp fragment. 20μl of the reaction mixture is electrophoresed on 1% agarose gels, stained with ethidium bromide and transferred to nylon membranes (Gene Screen). Filters were probed with a ^{32}P-labelled EcoRI-PstI about 1.4 kbp fragment located in the mouse heavy chain, 3' of the insertion site (see Chart 2). For further analysis, genomic DNA was prepared from ES cells, digested with restriction enzymes as recommended by the manufacturers, and fragments were separated on 1% agarose gels. DNA was transferred to nylon membranes (Gene Screen) and

probed with the ^{32}P-labelled fragment as described above.

Analysis of G418-resistant ES colonies

In the first experiment, PCR analysis of the pooled colonies detected one positive PCR signal of the expected size (about 1.4kbp) out of 34 pools representing 136 G418-resistant colonies. The four individual colonies that had contributed to this positive pool were analyzed individually by PCR, and a positive clone, ES33D5, was identified. Similar analysis of 540 G418-resistant colonies obtained in the second experiment yielded 4 additional positive clones (ES41-1, ES61-1, ES65-1, ES110-1).

In order to verify the targeting disruption of one copy of the J genes, (the gene is autosomal and thus present in two copies), the PCR positive clones were expanded and genomic DNA was prepared, digested with HindIII or with SacI and analyzed by Southern analysis as described using the EcoRI-PstI probe.

The replacement of the J genes by insertion of the neomycin gene by an homologous recombination event results in an HindIII fragment, detectable with the EcoRI-PstI probe, which is about 1.9 kbp longer than the equivalent fragment in the native locus, due to the loss of two HindIII sites located in the deleted J gene region (see Chart 2C). Southern analysis of each of the 5 positive clones by HindIII digestion gave a pattern which indicated that one of the two copies of the heavy chain J genes had been disrupted. Three labelled fragments were detected: one fragment (about 760 bp), identical in size to that present in untreated cells at the same intensity, one fragment (about 2.3kbp) identical in size to that present in untreated cells, but of decreased intensity in the PCR positive clone, and an additional fragment about 4.2 kbp, the size predicted for an homologous

recombination event, present only in the PCR-positive clones. Similarly, the replacement of the J genes by the neomycin gene by an homologous recombination event results in a loss of one SacI site and the appearance of a fragment, detectable with the EcoRI-PstI probe, which is about 570 bp smaller than the equivalent fragment in the native locus (see Chart 2C). Southern analysis of the clones by SacI digestion gave the expected pattern of one native and one targeted allele: about 4.0 kbp fragment, identical in size to that detected in untreated cells, but of decreased intensity in the 5 positive clones, and an additional fragment of about 3.4 kbp, the size predicted for a targeted homologous recombination event, present only in the identified clones. Rehybridization of the Southern blots with a probe for the neomycin gene shows that only the 4.2 kbp and 3.4 kbp fragments, resulting from the HindIII and the SacI digestion, respectively, hybridized to the probe as predicted by the targeting event.

Inactivation of mouse immunglobulin heavy chain J genes in mice

Injection of targeted ES cells into mouse blastocysts and generation of chimeric offsprings

Mice were purchased from Jackson Laboratories (Bar Harbor, ME). Three and a half day old C57BL/6 blastocysts were obtained from 4-5 week old superovulated females as described by Koller et al. 1989 (supra). ES cells were trypsinized, washed once with fresh DMEM media and diluted to about $1x10^6$/ml in M2 media. About 5μl of cells were added to a 150μl droplet of M2 media, under paraffin oil, containing the blastocysts. Ten to fifteen cells were injected into the blastocoel of each blastocyst. Six to nine ES cell-containing blastocysts were returned to each uterine horn of C57BL/6 x DBA F1

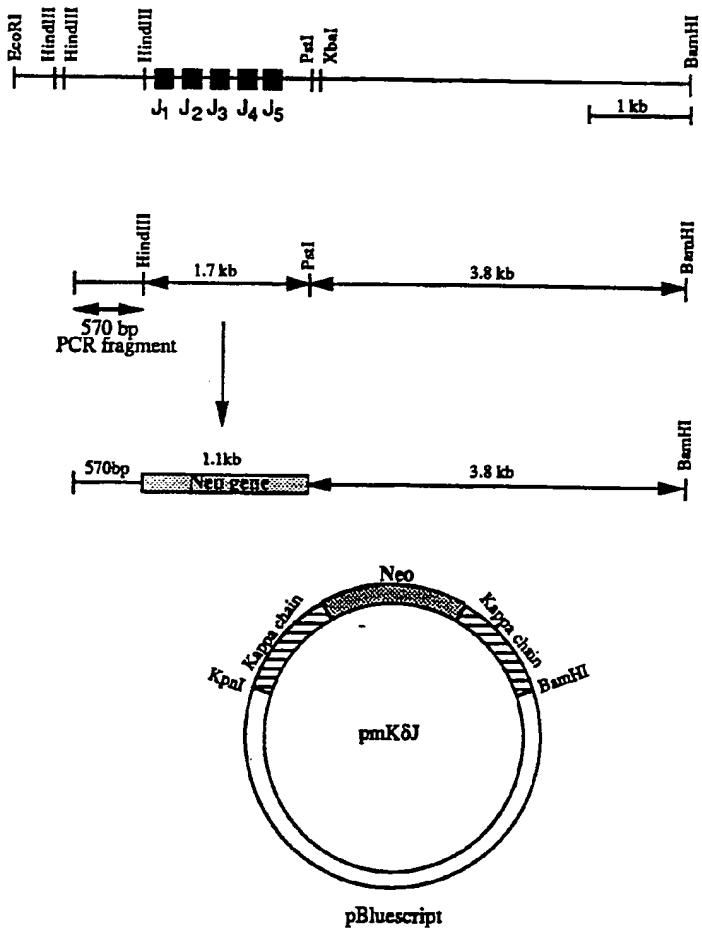
pseudopregnant females mated 2.5 days previously with vasectomized males. Pups derived from the injected blastocysts were generally born 16-18 days later. The contribution of the ES cells to the offspring was judged visually by examination of the coat color of the pups. The blastocysts were obtained from C57BL/6 mice, which are solid black in color. The ES cell line E14TG2a, the parental line from which the targeted cell lines were derived, was isolated from 129/01a mice. This mouse strain is cream in color, the combined effect of three color genes, the dominant A^{w} allele at the agouti locus, the recessive pink-eyed-dilute allele at the p locus and the recessive C^{ch} allele the C locus. Offspring in which the ES cells participated in the formation of the animal had coats containing brown and cream hair. The ES cell line ES41-1 carrying inactivated mouse immunoglobulin heavy chain, was injected into C57BL/6 mouse blastocysts as described above. Six out of the 18 surviving pups had a high degree of coat color chimerism (70-90%). PCR analysis of DNA isolated from chimeric newborn pups from a female implanted with blastocysts injected with the inactivated ES cells, indicated that the mutated immunoglobulin heavy chain locus is present in a variety of organs such as spleen, thymus, kidney, liver, brain and skin.

Inactivation of the mouse Ig kappa chain J genes in ES cells

A 5.6 Kb HindIII-BamHI fragment, containing the mouse immunoglobulin kappa chain J region genes and 3' flanking sequences, cloned from a Balb/c mouse embryo genomic library and inserted into pBluescriptSK vector to yield the plasmid (pKJ). pKJ was digested with HindIII and PstI to delete an about 1.7 Kb fragment containing the 5 J genes (see Chart 3). A 570 bp blunt HindIII fragment, spanning the

Chart 3

Kappa chain inactivation vector



region 5′ to the HindIII site adjacent to the kappa J region, was cloned from mouse genomic DNA by polymerase chain reaction (PCR). This fragment was inserted into HindIII-SmaI digested pIc cloning vector (Marsh et al., 1984, Gene, 32:481-485) and was excised by digestion with KpnI-XhoI. An about 1.1 Kb XhoI-BamHI fragment, blunted at the BamHI site, containing the neomycin resistance gene driven by the Herpes simlex virus thymidine kinase gene (HSV-tk) promoter and polyoma enhancer was isolated from pMC1Neo (Thomas and Capecchi, 1987, supra). The neomycin fragment was inserted into the HindIII-PstI deleted pKJ, which was blunted at the PstI site, 5′ to the kappa sequences. The resulting plasmid was digested with KpnI and XhoI and the 570 bp KpnI-XhoI kappa fragment was inserted into the KpnI-XhoI cleaved vector, 5′ to the neomycin gene, to generate the inactivation vector (pmKδJ, see chart 3). The transcriptional orientation of the neomycin and the kappa chain genes is the same in pmKδJ. The plasmid was lineraized by ApaLI before transfection into ES cells. The linearized sequence has about 3.8 Kb and 570 bp of homology to the cellular sequences, located 3′ and 5′ to the neomycin gene, respectively.

Analysis of G418-resistant ES colonies

Electroporation of the kappa inactivation vector into ES cells and screening for homologous recombination events was carried out as described for the inactivation of the immunoglobulin heavy chain. G418-resistant ES colonies were analysed for homologous recombination targeting by PCR using two priming oligonucleotides CGGTTGCTGTTGTATCCATAACTC and CATCAGAGCAGCCGATTGTCTG, which correspond respectively to the sequences located in the mouse kappa chain gene, about 67 bp 5′ of the insertion site, and about 370 bp 3′ of the XhoI site of the neomycin gene. A

^{32}P-labelled 80 base oligonucleotide, which starts about 10 bp 5′ of the insertion site, was used as a probe to detect the targeted PCR product. Successful homologous recombination gives rise to an about 1030 bp fragment. PCR analysis of 650 G418-resistant colonies detected 3 positive colonies (ES56-1, ES69-4, ES147-1). Southern analysis of these colonies confirmed the integration of the inactivation vector into one allele of the kappa immunoglobulin loci leading to a deletion of the J region.

Production of human Ig in transgenic mice

Example: production of human heavy chain in transgenic mice DNA vector

An SpeI fragment, spanning the human heavy chain VH6-D-J-Cμ-Cδ region (Berman et al., EMBO J. (1988) 7: 727-738; see Chart 4) is isolated from a human library cloned into a yeast artificial chromosome (YAC) vector (Burke, et al., Science, 236: 806-812) using DNA probes described by Berman et al. (EMBO J. (1988) 7:727-738). One clone is obtained which is estimated to be about 100 Kb. The isolated YAC clone is characterized by pulsed-field gel electrophoresis (Burke et al., supra; Brownstein et al., Science, 244: 1348-1351), using radiolabelled probes for the human heavy chain (Berman et al., supra).

Introduction of YAC clones into embryos

High molecular weight DNA is prepared in agarose plugs from yeast cells containing the YAC of interest (i.e., a YAC containing the aforementioned SpeI fragment from the IgH locus). The DNA is size-fractionated on a CHEF gel apparatus and the YAC band is cut out of the low melting point agarose gel. The gel fragment is equilibrated with polyamines and then melted and treated with agarase to digest the agarose.

The polyamine-coated DNA is then injected into the male pronucleus of fertilized mouse embryos which are surgically introduced into the uterus of a psueudopregnant female as described above. The transgenic nature of the newborns is analyzed by a slot-blot of DNA isolated from tails and the production of human heavy chain is analyzed by obtaining a small amount of serum and testing it for the presence of Ig chains with rabbit anti-human antibodies.

As an alternative to microinjection, YAC DNA is transferred into murine ES cells by ES cell: yeast protoplast fusion (Traver et al., 1989 Proc. Natl. Acad. Sci., USA, 86:5898-5902; Pachnis et al., 1990, ibid 87: 5109-5113). First, the neomycin-resistance gene from pMC1Neo and a yeast selectable marker are inserted into nonessential YAC vector sequences in a plasmid. This construct is used to transform a yeast strain containing the IgH YAC, and pMC1Neo is integrated into vector sequences of the IgH YAC by homologous recombination. The modified YAC is then transferred into an ES cell by protoplast fusion (Traver et al., 1989; Pachnis et al., 1990), and resulting G418-resistant ES cells which contain the intact human IgH sequences are used to generate chimeric mice.

Production of human Ig by chimeric mice

Construction of human heavy chain replacement vector.

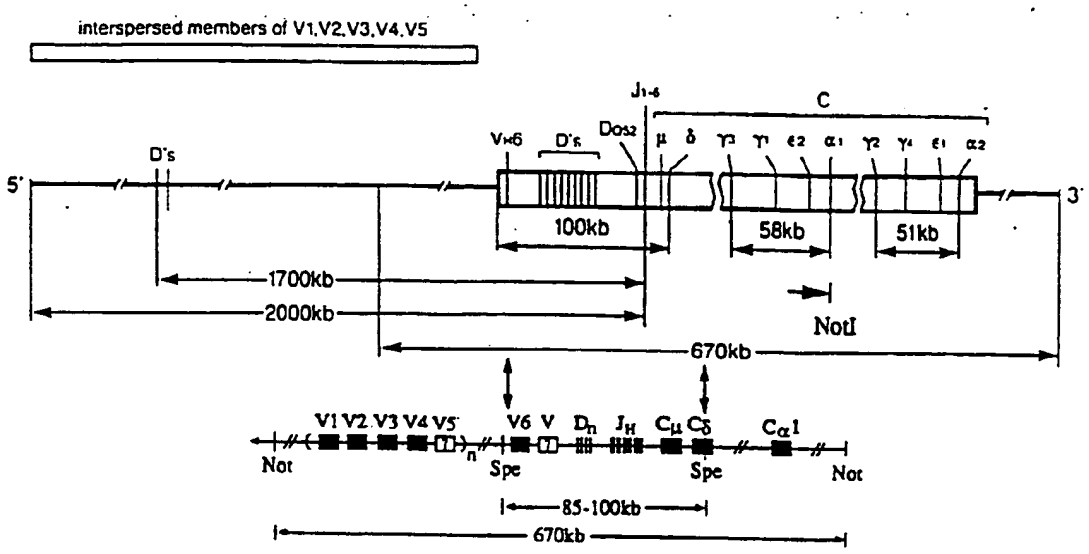
The replacing human sequences include the SpeI 100 kbp fragment of genomic DNA which encompasses the human VH6-D-J-Cμ-Cδ heavy chain region isolated from a human-YAC library as described before. The flanking mouse heavy chain sequences, which drive the homologous recombination replacement event, contain a 10 kbp BamHI fragment of the mouse Cϵ-Cα heavy chain and a 5' J558 fragment comprising the 5' half of the J558 fragment of the mouse heavy chain variable

region, at the 3' and 5' ends of the human sequences, respectively (Chart 4). These mouse sequences are isolated from a mouse embryo genomic library using the probes described in Tucker et al., PNAS USA, 78: 7684-7688, 1981, and Blankenstein and Krawinkel (1987, supra), respectively. The 1150 bp XhoI to BamHI fragment, containing a neomycin-resistance gene driven by the Herpes simplex virus thymidine kinase gene (HSV-tk) promoter and a polyoma enhancer is isolated from pMC1Neo (Koller and Smithies, 1989, supra). A synthetic adaptor is added onto this fragment to convert the XhoI end into a BamHI end and the resulting fragment is joined to the BamHI mouse Cε-Cα in a plasmid.

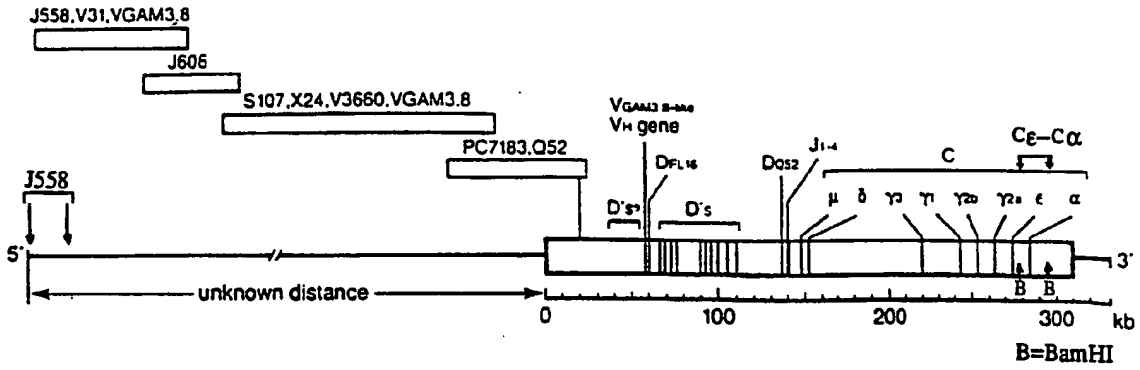
From the YAC clone containing the human heavy chain locus, DNA sequences from each end of the insert are recovered either by inverse PCR (Silverman et al., PNAS, 86:7485-7489, 1989), or by plasmid rescue in E. coli, (Burke et al., 1987; Garza et al. Science, 246:641-646, 1989; Traver et al., 1989) (see Chart 4). The isolated human sequence from the 5'V6 end of the YAC is ligated to the mouse J558 sequence in a plasmid and likewise, the human sequence derived from the 3'Cδ end of the YAC is ligated to the Neo gene in the plasmid containing Neo and mouse Cε-Cα described above. The human V6-mouse J558 segment is now subcloned into a half-YAC cloning vector that includes a yeast selectable marker (HIS3) not present in the original IgH YAC, a centromere (CEN) and a single telomere (TEL). The human Cδ - Neo - mouse Cε - Cα is likewise subcloned into a separate half-YAC vector with a different yeast selectable marker (LEU2) and a single TEL. The half-YAC vector containing the human V6 DNA is linearized and used to transform a yeast strain that is deleted for the chromosomal HIS3 and LEU2 loci and which carries the IgH YAC. Selection for histidine-prototrophy gives rise to

Chart 4

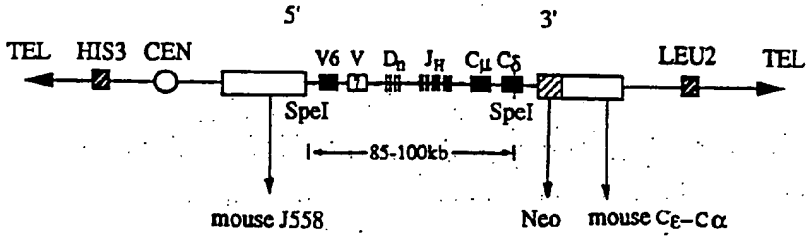
(A) Human heavy chain locus



(B) Mouse heavy chain locus



(C) Human heavy chain replacement YAC vector



yeast colonies that have undegone homoloous recombination between the human V6 DNA sequences and contain a recombinant YAC. The half-YAC vector containing the human Cδ DNA is then linearized and used to transform the yeast strain generated in the previous step. Selection for leucine-prototrophy results in a yeast strain containing the complete IgH replacement YAC (see Chart 4). This YAC is isolated and introduced into ES cells by microinjection as described previously for embryos.

In accordance with the above procedures, an antigenic or chimeric non-primate host, particularly a mouse host, may be produced which can be immunized to produce human antibodies or analogs specific for an immunogen. In this manner, the problems associated with obtaining human monoclonal antibodies are avoided, since mice can be immunized with immunogens which could not be used with a human host. Furthermore, one can provide for booster injections and adjuvants which would not be permitted with a human host. The resulting B-cells may then be used for immortalization for the continuous production of the desired antibody. The immortalized cells may be used for isolation of the genes encoding the immunoglobulin or analog and be subjected to mutation by in-vitro mutagenesis or other mutagenizing technique to modify the properties of the antibodies. These mutagenized genes may then be returned to the immortalized cells for homologous recombination to provide for a continuous mammalian cellular source of the desired antibodies. The subject invention provides for a convenient source of human antibodies, where the human antibodies are produced in analogous manner to the production of antibodies in a human host. The mouse cells conveniently provide for the activation and rearrangement of human DNA in mouse cells for production of human antibodies.

All publications and patent applications cited in this specification are herein incorporated by reference as if each individual publication or patent application were specifically and individually indicated to be incorporated by reference.

Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it will be readily apparent to those of ordinary skill in the art in light of the teachings of this invention that certain changes and modifications may be made thereto without departing from the spirit or scope of the appended claims.

WHAT IS CLAIMED IS:

1. A method for producing xenogeneic primate antisera or antibody analog in a mammalian non-primate host, said method comprising:

immunizing said host with an immunogen under conditions to stimulate an immune response to said immunogen, whereby said host mounts an immune response to said immunogen and produces plasma cells, B-cells having surface immunoglobulin specific for said immunogen and antibodies specific for said immunogen,

wherein said non-primate host is characterized by being substantially incapable of expressing, (1) an endogenous immunoglobulin heavy chain; and, (2) at least one endogenous light chain; and capable of expressing an exogenous primate heavy chain and light chain or analog.

2. A method according to Claim 1, wherein said host is a mouse.

3. A method according to Claim 1, wherein said analog comprises a variable region joined by a peptide bond to a peptide other than solely the immunoglobulin constant region.

4. A method according to Claim 1, wherein said antisera is human antisera.

5. A method according to Claim 1, including the additional step of immortalizing said B-cells for production of monoclonal antibodies.

6. A method according to Claim 1, wherein said host comprises B-cells comprising a functional immunoglobulin locus comprising an exogenous variable region and at least one human constant region.

7. A method according to Claim 1, wherein said host comprises at least a functional portion of human immunoglobulin loci for a heavy chain and a light chain and inactivated endogenous loci for the two copies of the heavy chain and the two copies of at least one light chain type, so as to be substantially incapable of expressing an endogenous antibody subunit.

8. A transgenic non-primate mammal comprising a genome comprising:

at least a functional portion of human immunoglobulin loci for at least a portion of the heavy chain domain and a light chain domain and inactivated endogenous loci for the two copies of the heavy chain and the two copies of at least one light chain type, so as to be substantially incapable of expressing an endogenous antibody subunit.

9. A transgenic non-primate mammal according to Claim 8, wherein said human immunoglobulin loci are at the analogous endogenous loci.

10. A transgenic non-primate mammal according to Claim 8, wherein said human immunoglobulin loci are at other than the analogous endogenous loci.

11. A transgenic non-primate mammal according to Claim 8, wherein said mammal is murine.

12. A transgenic non-primate mammal according to Claim 8, wherein said at least a portion of the heavy chain domain comprises at least one constant type region.

13. A transgenic non-primate mammal according to Claim 8, wherein said at least a portion of the heavy chain domain comprises a peptide sequence other than a constant region.

14. A transgenic mouse comprising a genome comprising:

inactivated loci of both copies of the immunoglobulin heavy chain and both copies of at least one type of light chain as a result of at least one of a deletion or insertion at the loci; and

at least a functional portion of human immunoglobulin loci for a heavy chain and a light chain.

15. A transgenic mouse according to Claim 14, wherein said at least a functional portion comprises only a portion of the variable region.

16. A transgenic mouse comprising a genome comprising:

inactivated loci of both copies of the immunoglobulin heavy chain and both copies of at least one type of light chain as a result of at least one of a deletion or insertion at the loci; and

at least a functional portion of human immunoglobulin loci for a heavy chain and a light chain, wherein the constant region of at least one of said heavy chain and said light chain comprises a sequence other than said constant region sequence.

17. A transgenic non-primate mammal comprising a genome comprising:

at least a functional portion of human immunoglobulin loci for at least a portion of the heavy chain and a light chain and inactivated endogenous loci as a result of homologous recombination at said loci for the two copies of the

heavy chain and the two copies of at least one light chain type, so as to be substantially incapable of expressing an endogenous antibody subunit.

18. A transgenic non-primate mammal according to Claim 17, wherein said mammal is a mouse and said human immunoglobulin heavy and light chain loci are at mouse immunoglobulin heavy and light chain loci, respectively.

19. A non-primate embryonic stem cell comprising a lesion in the J region of one immunoglobulin locus resulting in the incapacity to rearrange to a functional subunit of an immunoglobulin.

20. A stem cell according to Claim 19, wherein said subunit is the heavy chain.

21. A stem cell according to Claim 20, wherein said non-primate is a mouse.

22. A stem cell according to Claim 19, wherein said lesion is an insertion.

23. A stem cell according to Claim 22, wherein said insertion is a marker capable of selection.

24. A stem cell according to Claim 19 wherein said lesion is a deletion.

25. A stem cell according to Claim 19, wherein said non-primate is a mouse.

INTERNATIONAL SEARCH REPORT

International Application PCT/US 91/00245

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) [5]

According to International Patent Classification (IPC) or to both National Classification and IPC

see attached sheet

II. FIELDS SEARCHED

Minimum Documentation Searched [6]

Classification System	Classification Symbols
see attached sheet	

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched [7]

see attached sheet

III. DOCUMENTS CONSIDERED TO BE RELEVANT [14]

Category *	Citation of Document, [16] with indication, where appropriate, of the relevant passages [17]	Relevant to Claim No [18]
see attached sheet		

* Special categories of cited documents: [15]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search [2]	Date of Mailing of this International Search Report [3]
14 MARCH 1991	26 APR 1991
International Searching Authority [1]	**Signature of Authorized Officer [20]**
ISA/US	GABRIELLE E. BUGAISKY

Form PCT/ISA/210 (second sheet) (May 1986)

PCT/US/ 91/00245
Attachment to Form PCT/ISA/210
I. Classification of subject matter
IPC(5): C12P 21/06; C12N 15/00
U.S. Cl.: 435/69.1, 172.3; 800/2

II. Fields searched
U.S. Cl. 435/69.1, 69.6, 70.1, 172.3; 436/547; 530/387; 800/2; 935/22, 65, 106

Databases: Dialog Information Services Online (File sets Medline and World Patent Index)
Automated Patent System (File USPAT)

gene transfer or gene replacement or gene inactivation, homologous recombination; embryonic stem cell, animal stem cell, embryonal carcinoma, transgenic animal or mammal, xenogeneic antibody or antiserum or immune response, immunoglobulin; immunoglobulin gene.

III. Documents considered relevant

Category	Citation	Claims
Y,P	US, A. 4,959,313 (TAKETO) 25 September, 1990 see entire document.	19-25
Y,P	US, A. 4,950,599 (BERTLING) 21 August, 1990 see entire document.	8-25
Y	Proc. Natl. Acad. Sci., USA. Vol. 83, issued April 1986, K.-I. Yamamura, et al., "Cell-type-specific and regulated expression of a human γ1 heavy-chain immunoglobulin gene in transgenic mice", pages 2152-2156, see entire document.	1-25
Y	Proc. Natl. Acad. Sci., USA. Vol. 86, issued November 1989, B. Koller, et al., "Inactivating the β2-microglobulin gene in mouse embryonic stem cells by homologous recombination", pages 8932-8935, see entire document.	1-25
A	Proc. Natl. Acad. Sci., USA. Vol. 83, issued July 1966, D. Ayares, et al., "Sequence homology requirements for intermolecular recombination in mammalian cells", pages 5199-5203, see entire document.	8-25
A	Proc. Natl. Acad. Sci., USA. Vol. 85, issued February 1988, R.Brinster, et al., "Introns increase transcriptional efficiency in transgenic mice", pages 836-840, see entire document.	1-25
Y	Prog. Nucleic Acid Res. Mol. Biol., Vol 36, issued 1989, R. Kucherlapalati, "Homologous recombination in mammalian somatic cells", pages 301-310, see entire document.	1-25
Y	Proc. Natl. Acad. Sci., USA. Vol. 86, issued October 1989, A. Shimizu, et al., "Immunoglobulin double-isotype expression by trans-mRNA in a human immunoglobulin transgenic mouse", pages 8020-8023, see entire document.	1-25

Attachment to Telephone Memorandum
PCT/US91/00245

Observations where unity of invention is lacking

Detailed reasons for holding lack of Unity of Invention.

There are three groups of claims: Group I is a method for producing antisera; transgenic animals; Group II is for embryonic stem cells. Group I is related as first mentioned product and process of use. Group II consists of a second mentioned product, which can exist independently of the first mentioned product. PCT Rules 13.1 and 13.2 do not provide for multiple products.

Itemized summary of claims groupings

I. Claims 1-7, drawn to a method for producing xenogeneic antisera, classified in Class 435, subclass 69.1.

Claims 8-18, drawn to transgenic animals with lesions in endogenous immunoglobulin genes, so that they can only express human immunoglobulin genes, classified in Class 800, subclass 2.

II. Claims 19-25, drawn to embryonic stem cells with lesions in endogenous immunoglobulin genes, classified in Class 435, subclass 230.1.

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification [5]: C12N 15/13, C07K 15/28	A1	(11) International Publication Number: WO 91/10737
		(43) International Publication Date: 25 July 1991 (25.07.91)

(21) International Application Number: PCT/US91/00209

(22) International Filing Date: 10 January 1991 (10.01.91)

(30) Priority data:
464,350 11 January 1990 (11.01.90) US

(71) Applicant: MOLECULAR AFFINITIES CORPORATION [US/US]; 291 Broadway, New York, NY 10007 (US).

(72) Inventor: WIGLER, Michael, H. ; 1 Walden Court, Lloyd Harbor, NY 11743 (US).

(74) Agents: GRANAHAN, Patricia et al.; Hamilton, Brook, Smith & Reynolds, Two Militia Drive, Lexington, MA 02173 (US).

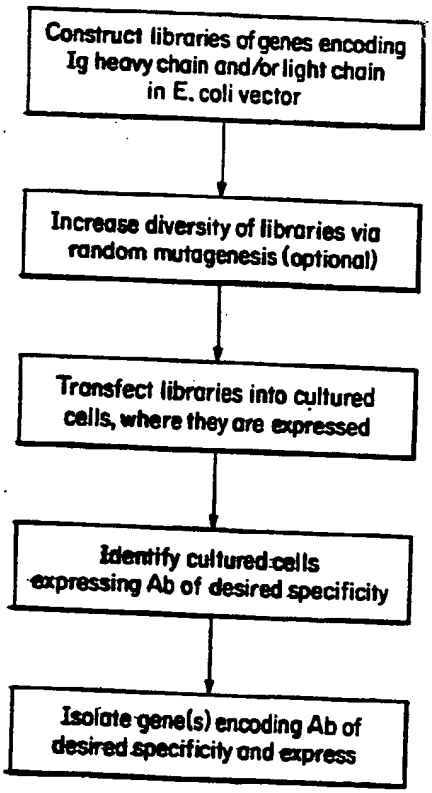
(81) Designated States: AT, AT (European patent), AU, BB, BE (European patent), BF (OAPI patent), BG, BJ (OAPI patent), BR, CA, CF (OAPI patent), CG (OAPI patent), CH, CH (European patent), CM (OAPI patent), DE, DE (European patent), DK, DK (European patent), ES, ES (European patent), FI, FR (European patent), GA (OAPI patent), GB, GB (European patent), GR, GR (European patent), HU, IT (European patent), JP, KP, KR, LK, LU, LU (European patent), MC, MG, ML (OAPI patent), MR (OAPI patent), MW, NL, NL (European patent), NO, RO, SD, SE, SE (European patent), SN (OAPI patent), SU, TD (OAPI patent), TG (OAPI patent).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: PRODUCTION OF ANTIBODIES USING GENE LIBRARIES

(57) Abstract

A method of producing libraries of genes encoding antigen-combining molecules or antibodies; a method of producing antigen-combining molecules which does not require an *in vivo* procedure; a method of obtaining antigen-combining molecules of selected specificity which does not require an *in vivo* procedure; vectors useful in the present method; and antigen-combining molecules produced by the method. The antigen-combining molecules are useful for the detection, quantitation, purification and neutralization of antigens, as well as for diagnostic, therapeutic and prophylactic purposes.



Attorney Docket No: 10271-054-999
Application Serial No.: 10/091,268
Reference B38

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**ES**	Spain	**MG**	Madagascar
AU	Australia	**FI**	Finland	**ML**	Mali
BB	Barbados	**FR**	France	**MN**	Mongolia
BE	Belgium	**GA**	Gabon	**MR**	Mauritania
BF	Burkina Faso	**GB**	United Kingdom	**MW**	Malawi
BG	Bulgaria	**GN**	Guinea	**NL**	Netherlands
BJ	Benin	**GR**	Greece	**NO**	Norway
BR	Brazil	**HU**	Hungary	**PL**	Poland
CA	Canada	**IT**	Italy	**RO**	Romania
CF	Central African Republic	**JP**	Japan	**SD**	Sudan
CG	Congo	**KP**	Democratic People's Republic of Korea	**SE**	Sweden
CH	Switzerland			**SN**	Senegal
CI	Côte d'Ivoire	**KR**	Republic of Korea	**SU**	Soviet Union
CM	Cameroon	**LI**	Liechtenstein	**TD**	Chad
CS	Czechoslovakia	**LK**	Sri Lanka	**TG**	Togo
DE	Germany	**LU**	Luxembourg	**US**	United States of America
DK	Denmark	**MC**	Monaco		

Production of antibodies using gene libraries.

Description

Background of the Invention

Monoclonal and polyclonal antibodies are useful for a variety of purposes. The precise antigen specificity of antibodies makes them powerful tools that can be used for the detection, quantitation, purification and neutralization of antigens.

Polyclonal antibodies are produced in vivo by immunizing animals, such as rabbits and goats, with antigens, bleeding the animals and isolating polyclonal antibody molecules from the blood. Monoclonal antibodies are produced by hybridoma cells, which are made by fusing, in vitro, immortal plasmacytoma cells with antibody producing cells (Kohler, G. and C. Milstein, Nature, 256:495 (1975)) obtained from animals immunized in vivo with antigen.

Current methods for producing polyclonal and monoclonal antibodies are limited by several factors. First, methods for producing either polyclonal or monoclonal antibodies require an in vivo immunization step. This can be time consuming and require large amounts of antigen. Second, the repertoire of antibodies expressed in vivo is restricted by physiological processes, such as those which mediate self-tolerance that disable autoreactive B cells (Goodnow, C.C., et al., Nature, 334:676

(1988); Goodnow, J.W., Basic and Clinical Immunology, Ed. 5, Los Altos, CA, Large Medical Publications (1984); Young, C.R., Molecular Immunology, New York, Marcel Dekker (1984)). Third, although antibodies can exist in millions of different forms, each with its own unique binding site for antigen, antibody diversity is restricted by genetic mechanisms for generating antibody diversity (Honjo, T., Ann. Rev. Immunol., 1:499 (1983); Tonegawa, S., Nature:302:575 (1983)). Fourth, not all the antibody molecules which can be generated will be generated in a given animal. As a result, raising high affinity antibodies to a given antigen can be very time consuming and can often fail. Fifth, the production of human antibodies of desired specificity is very problematical.

A method of producing antibodies which avoids the limitations of presently-available methods, such as the requirement for immunization of an animal and in vivo steps, would be very useful, particularly if it made it possible to produce a wider range of antibody types than can be made using presently-available techniques and if it made it possible to produce human antibody types.

Disclosure of the Invention

The present invention relates to a method of producing libraries of genes encoding antigen-combining molecules or antibodies; a method of producing antigen-combining molecules, also referred to as antibodies, which does not require an in vivo procedure, as is required by presently-available methods; a method of obtaining antigen-combining molecules (antibodies) of

selected or defined specificity which does not require an in vivo procedure; vectors useful in the present method and antibodies produced or obtained by the method.

The present invention relates to an in vitro process for synthesizing DNA encoding families of antigen-combining molecules or proteins. In this process, DNA containing genes encoding antigen-combining molecules is obtained and combined with oligonucleotides which are homologous to regions of the genes which are conserved. Sequence-specific gene amplification is then carried out using the DNA containing genes encoding antigen-combining proteins as template and the homologous oligonucleotides as primers.

This invention also relates to a method of creating diverse libraries of DNAs encoding families of antigen-combining proteins by cloning the product of the in vitro process for synthesizing DNA, described in the preceeding paragraph, into an appropriate vector (e.g., a plasmid, viral or retroviral vector).

The subject invention provides an alternative method for the production of antigen-combining molecules, which are useful affinity reagents for the detection and neutralization of antigens and the delivery of molecules to antigenic sites. The claimed method differs from production of polyclonal antibody molecules derived by immunization of live animals and from production of monoclonal antibody molecules through the use of hybridoma cell lines in that it does not require an in vivo immunization step, as do presently available methods. Rather, diverse libraries of genes which encode antigen-combining sites comprising a significant proportion of an

animal's repertoire of antibody combining sites are made, as described in detail herein. These genes are expressed in living cells, from which molecules of desired antigenic selectivity can be isolated and purified for various uses.

Antigen-combining molecules are produced by the present method in the following manner, which is described in greater deail below. Initially, a library of antibody genes which includes a set of variable regions encoding a large, diverse and random group of specificities derived from animal or human immunoglobulins is produced by amplifying or cloning diverse genomic fragments or cDNAs of antibody mRNAs found in antibody-producing tissue.

In an optional step, the diversity of the resulting libraries can be increased by means of random mutagenesis. The gene libraries are introduced into cultured host cells, which may be eukaryotic or prokaryotic, in which they are expressed. Genes encoding antibodies of desired antigenic specificity are identified, using a method described herein or known techniques, isolated and expressed in quantities in appropriate host cells, from which the encoded antibody can be purified.

Specifically, a library of genes encoding immunoglobulin heavy chain regions and a library of genes encoding immunoglobulin light chain regions are constructed. This is carried out by obtaining antibody-encoding DNA, which is either genomic fragments or cDNAs of antibody mRNAs, amplfying or cloning the fragments or cDNAs; and introducing them into a standard framework antibody gene vector, which is used to introduce the

antibody-encoding DNA into cells in which the DNA is expressed. The vector includes a framework gene encoding a protein, such as a gene encoding an antibody heavy chain or an antibody light chain which can be of any origin (human, non-human) and can be derived from any of a number of existing DNAs encoding heavy chain immunoglobulins or light chain immunoglobulins. Such vectors are also a subject of the present invention and are described in greater detail in a subsequent section. Genes from one or both of the libraries are introduced into appropriate host cells, in which the genes are expressed, resulting in production of a wide variety of antigen-combining molecules.

Genes encoding antigen-combining molecules of desired specificity are identified by identifying cells producing antigen-combining molecules which react with a selected antigen and then obtaining the genes of interest. The genes of interest can subsequently be introduced into an appropriate host cell (or can be further modified and then introduced into an appropriate host cell) for further production of antigen-combining molecules, which can be purified and used for the same purposes for which conventionally-produced antibodies are used.

Through use of the method described, it is possible to produce antigen-combining molecules which are of wider diversity than are antibodies available as a result of known methods; novel antigen-combining molecules with a diverse range of specificities and affinities and antigen-combining molecules which are predominantly human in origin. Such antigen-combining molecules are a

subject of the present invention and can be used clinically for diagnostic, therapeutic and prophylactic purposes, as well as in research contexts, and for other purposes.

Brief Description of the Drawings

Figure 1 is a schematic representation of the method of the present invention by which antigen-combining molecules, or antibodies, are produced.

Figure 2 is a schematic representation of amplification or cloning of IgM heavy chain variable region DNA from mRNA, using the polymerase chain reaction.
Panel A shows the relevant regions of the poly adenylated mRNA encoding the secreted form of the IgM heavy chain. S denotes the sequences encoding the signal peptide which causes the nascent peptide to cross the plasma membrane. V, D and J together comprise the variable region. C_H1, C_H2, and C_H3 are the three constant domains of $C\mu$. Hinge encodes the hinge region. C, B and Z are oligonucleotide PCR primers (discussed below).

Panel B shows the reverse transcript DNA product of the mRNA primed by oligonucleotide Z, with the addition of poly-dC by terminal transferase at the 3' end.
Panel C is a schematic representation of the annealing of primer A to the reverse transcript DNA.
Panel D shows the final double stranded DNA PCR product made utilizing primers A and B.
Panel E shows the product of PCR annealed to primer C.
Panel F is a blowup of Panel E, showing in greater detail the structure of primer C. Primer C consists of two

parts: a 3' part complementary to IgM heavy chain mRNA as shown, and a 5' part which contains restriction site RE2 and spacer.

Panel G shows the final double stranded DNA PCR product made utilizing primers A and C and the product of the previous PCR (depicted in D) as template. The S, V, D, J regions are again depicted.

Figure 3 is a schematic representation of the heavy chain framework vector pFHC. The circular plasmid (above) is depicted linearized (below) and its relevant components are shown: animal cell antibiotic resistance marker; bacterial replication origin; bacterial cell antibiotic resistance marker; Cμ enhancer; LTR containing the viral promoter from the Moloney MLV retrovirus DNA; PCR primer (D); cDNA cloning site containing restriction endonuclease sites, RE1 and RE2, separated by spacer DNA; Cμ exons; and poly A addition and termination sequences derived from the Cμ gene or having the same sequence as the Cμ gene.

Figure 4 depicts a nucleotide sequence of the C_H1 exon of the Cμ gene, and its encoded amino acid sequence (Panel A). The nucleotide coordinate numbers are listed above the line of nucleotide sequences. Panel B depicts the N-doped sequence, as defined in the text.

Detailed Description of the Invention

The present invention provides a method of producing antigen-combining molecules (or antibodies) which does not require an in vivo immunization procedure and which makes it possible to produce antigen-combining molecules

with far greater diversity than is shown by antibodies produced by currently-available techniques.

The present invention relates to a method of producing libraries of genes encoding antigen-combining molecules (antibody proteins) with diverse antigen-combining specificities; a method of producing such antigen-combining molecules, antigen-combining molecules produced by the method and vectors useful in the method. The following is a description of generation of such libraries, of the present method of producing antigen-combining molecules of selected specificity and of vectors useful in producing antigen-combining molecules of the present invention.

As described below, the process makes use of techniques which are known to those of skill in the art and can be applied as described herein to produce and identify antigen-combining molecules of desired antigenic specificity: the polymerase chain reaction (PCR), to amplify and clone diverse cDNAs encoding antibody mRNAs found in antibody-producing tissue; mutagenesis protocols to further increase the diversity of these cDNAs; gene transfer protocols to introduce antibody genes into cultured (prokaryotic and eukaryotic) cells for the purpose of expressing them; and screening protocols to detect genes encoding antibodies of the desired antigenic specificity. A general outline of the present method is represented in Figure 1.

Construction of Library of Genes Encoding Antigen-Combining Molecules

A key step in the production of antigen-combining molecules by the present method is the construction of a "library" of antibody genes which include "variable" regions encoding a large, diverse, but random set of specificities. The library can be of human or non-human origin and is constructed as follows:

Initially, genomic DNA encoding antibodies or cDNAs of antibody mRNA (referred to as antibody-encoding DNA) is obtained. This DNA can be obtained from any source of antibody-producing cells, such as spleen cells, peripheral blood cells, lymph nodes, inflammatory tissue cells and bone marrow cells. It can also be obtained from a genomic library or cDNA library of B cells. The antibody-producing cells can be of human or non-human origin; genomic DNA or mRNA can be obtained directly from the tissue (i.e., without previous treatment to remove cells which do not produce antibody) or can be obtained after the tissue has been treated to increase concentration of antibody-producing cells or to select a particular type(s) of antibody-producing cells (i.e., treated to enrich the content of antibody-producing cells). Antibody-producing cells can be stimulated by an agent which stimulates antibody mRNA production (e.g., lipopolysaccharide) before DNA is obtained.

Antibody-encoding DNA is amplified and cloned using a known technique, such as the PCR using appropriately-selected primers, in order to produce sufficient quantities of the DNA and to modify the DNA in such a manner

(e.g., by addition of appropriate restriction sites) that it can be introduced as an insert into an E. coli cloning vector. This cloning vector can serve as the expression vector or the inserts can later be introduced into an expression vector, such as the framework antibody gene vector described below. Amplified and cloned DNA can be further diversified, using mutagenesis, such as PCR, in order to produce a greater diversity or wider repertoire of antigen-binding molecules, as well as novel antigen-binding molecules.

Cloned antibody-encoding DNA is introduced into an expression vector, such as the framework antibody gene vector of the present invention, which can be a plasmid, viral or retroviral vector. Cloned antibody-encoding DNA is inserted into the vector in such a manner that the cloned DNA will be expressed as protein in appropriate host cells. It is essential that the expression vector used make it possible for the DNA insert to be expressed as a protein in the host cell. One expression vector useful in the present method is referred to as the framework antibody gene vector. Vectors useful in the present method contain antibody constant region or portions thereof in such a manner that when amplified DNA is inserted, the vector expresses a chimeric gene product comprising a variable region and a constant region in proper register. The two regions present in the chimeric gene product can be from the same type of immunoglobulin molecule or from two different types of immunoglobulin molecules.

These libraries of antibody-encoding genes are then expressed in cultured cells, which can be eukaryotic or

prokaryotic. The libraries can be introduced into host cells separately or together. Introduction of the antibody-encoding DNA *in vitro* into host cells (by infection, transformation or transfection) is carried out using known techniques, such as electroporation, protoplast fusion or calcium phosphate co-precipitation. If only one library is introduced into a host cell, the host cell will generally be one which makes the other antibody chain, thus making it possible to produce complete/functional antigen-binding molecules. For example, if a heavy chain library produced by the present method is introduced into host cells, the host cells will generally be cultured cells, such as myeloma cells or *E. coli*, which naturally produce the other (i.e., light) chain of the immunoglobulin or are engineered to do so. Alternatively, both libraries can be introduced into appropriate host cells, either simultaneously or sequentially.

Host cells in which the antibody-encoding DNA is expressed can be eukaryotic or prokaryotic. They can be immortalized cultured animal cells, such as a myeloma cell line which has been shown to efficiently express and secrete introduced immunoglobulin genes (Morrison, S.L., *et al.*, *Ann. N.Y. Acad. Sci.*, 507:187 (1987); Kohler, G. and C. Milstein, *Eur. J. Immunol.*, 6:511 (1976); Oi, V.T., *et al.*, *Immunoglobulin Gene Expression in Transformed Lymphoid Cells*, 80:825 (1983); Davis, A.C. and M.J. Shulman, *Immunol. Today*, 10:119 (1989)). One host cell which can be used to express the antibody-encoding DNA is the J558L cell line or the SP2/0 cell line.

Cells expressing antigen-combining molecules with a desired specificity for a given antigen can then be selected by a variety of means, such as testing for reactivity with a selected antigen using nitrocellulose layering. The antibodies identified thereby can be of human origin, nonhuman origin or a combination of both. That is, all or some of the components (e.g., heavy chain, light chain, variable regions, constant regions) can be encoded by DNA of human or nonhuman origin, which, when expressed produces the encoded chimeric protein which, in turn, may be human, nonhuman or a combination of both. In such antigen-combining molecules, all or some of the regions (e.g., heavy and light chain variable and constant regions) are referred to as being of human origin or of nonhuman origin, based on the source of the DNA encoding the antigen-combining molecule region in question. For example, in the case in which DNA encoding mouse heavy chain variable region is expressed in host cells, the resulting antigen-combining molecule has a heavy chain variable region of mouse origin. Antibodies produced may be used for such purposes as drug delivery, tumor imaging and other therapeutic, diagnostic and prophylactic uses.

Once antibodies of a desired binding specificity are obtained, their genes may be isolated and further mutagenized to create additional antigen combining diversity or antibodies of higher affinity for antigen.

Construction of Immunoglobulin Heavy Chain Gene Library and Production of Encoded Antigen-Binding Molecules

The following is a detailed description of a specific experimental protocol which embodies the concepts described above. Although the following is a description of one particular embodiment, the same procedures can be used to produce libraries in which the immunoglobulin and the heavy chain class are different or in which light chain genes are amplified and cloned. The present invention is not intended to be limited to this example. In the embodiment presented below, a diverse heavy chain gene library is constructed. Using the principles described in relation to the heavy chain gene library, a diverse light chain gene library is also constructed. These are co-expressed in an immortal tumor cell capable of producing antibodies, such as plasmacytoma cells or myeloma cells. Cells expressing antibody reactive to antigen are identified by a nitrocellulose filter overlay and antibody is prepared from cells identified as expressing it. As described in a subsequent section, there are alternative methods of library construction, other expression systems which can be used, and alternative selection systems for identifying antibody-producing cells or viruses.

Step 1 in this specific protocol is construction of libraries of genes in E. coli which encode immunoglobulin heavy chains. This is followed by the use of random mutagenesis to increase the diversity of the library, which is an optional procedure. Step 2 is introduction of the library, by transfection, into myeloma cells.

Step 3 is identification of myeloma cells expressing antibody with the desired specificity, using the nitrocellulose filter overlay technique or techniques known to those of skill in the art. Step 4 is isolation of the gene(s) encoding the antibody with the desired specificity and their expression in appropriate host cells, to produce antigen-combining fragments useful for a variety of purposes.

Construction

One key step in construction of the library of cDNAs encoding the variable region of mouse heavy chain genes is construction of an E. coli plasmid vector, designated pFHC. pFHC contains a "framework" gene, which can be any antibody heavy chain and serves as a site into which the amplified cloned gene product (genomic DNA or cDNA of antibody mRNAs) is introduced. pFHC is useful as a vector for this purpose because it contains RE1 and RE2 cloning sites. Other vectors which include a framework gene and other cloning sites can be used for this purpose as well. The framework gene includes a transcriptional promoter (e.g., a powerful promoter, such as a Moloney LTR (Mulligan, R.C., In Experimental Manipulation of Gene Expression, New York Adacemic Press, p. 155 (1983)) and a Cμ chain transcriptional enhancer to increase the level of transcriptions from the promoter (Gillies, S.D., et al., Cell, 33:717 (1983), a cloning site containing RE1 and RE2; part of the Cμ heavy chain gene encoding secreted protein; and poly A addition and termination sequences (Figure 3). The framework antibody gene vector of the present invention (pFHC) also includes a

selectable marker (e.g., an antibiotic resistance gene such as the neomycin resistance gene, neo^R) for animal cells; sequences for bacterial replication (ori); and a selectable marker (e.g., the ampicillin resistance gene, Amp^R) for bacterial cells. The framework gene can be of any origin (human, non-human), and can derive from any one of a number of existing DNAs encoding heavy chain immunoglobulins (Tucker, P.W., et al., Science, 206:1299 (1979); Honjo, T., et al., Cell, 18:559 (1979); Bothwell, A.L.M., et al., Cell, 24:625 (1981); Liu, A.Y, et al., Gene, 54:33 (1987); Kawakami, T., et al., Nuc. Acids. Res., 8:3933 (1980)). In this embodiment, the vector retains the introns between the C_H1, hinge, C_H2 and C_H3 exons. The "variable region" of the gene, which includes the V, D and J regions of the antibody heavy chain and which encodes the antigen binding site, is deleted and replaced with two consecutive restriction endonuclease cloning sites, RE1 and RE2. The restriction endonuclease site RE1 occurs just 3' to the LTR promoter and the restriction endonuclease site RE2 occurs within the constant region just 3' to the J region (see Figure 3).

Another key step in the production of antigen-combining molecules in this embodiment of the present invention is construction in an E. coli vector of a library of cDNAs encoding the variable region of mouse immunoglobulin genes. In this embodiment, the pFHC vector, which includes cloning sites designated RE1 and RE2, is used for cloning heavy chain variable regions, although any cloning vector with cloning sites having the same or similar characteristics (described below) can be used. Similarly, a light chain vector can be designed,

using the above described procedures and procedures known to a person of ordinary skill in the art.

In this embodiment, non-immune mouse spleens are used as the starting material. mRNA is prepared directly from the spleen or from spleen processed in such a manner that it is enriched for resting B cells. Enrichment of tissue results in a more uniform representation of antibody diversity in the starting materials. Lymphocytes can be purified from spleen using ficoll gradients (Boyum, A., Scand. J. of Clinical Invest., 21:77 (1968)). B cells are separated from other cells (e.g., T cells) by panning with anti-IgM coated dishes (Wysocki, L.J. and V.L. Sato, Proc. Natl. Acad. Sci., 75:2844 (1978)). Because activated cells express the IL-2 receptor but resting B cells do not, resting B cells can be separated yet further from activated cells by panning. Further purification by size fractionation on a Cell Sorter results in a fairly homogeneous population of resting B cells.

Poly A+ mRNA from total mouse spleen is prepared according to published methods (Sambrook, J., et al., Molecular Cloning: A Laboratory Manual, 2d Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1989)). Production of antibody mRNA can first be stimulated by lipopolysaccharide (LPS) (Andersson, J.A., et al., J. Exp. Med., 145:1511 (1977)). First strand cDNA is prepared to this mRNA population using as primer an oligonucleotide, Z, which is complementary to Cμ in the C_H1 region 3' to J. This primer is designated Z in Figure 2. First strand cDNA is then elongated by the terminal transferase reaction with dCTP to form a poly dC

tail (Sambrook, J., et al., Molecular Cloning: A Laboratory Manual, 2d Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1989)).

This DNA product is then used as template in a polymerase chain reaction (PCR) to amplify cDNAs encoding antibody variable regions (Saiki, R.K., et al., Science, 239:487 (1988); Ohara, O., et al., Proc. Natl. Acad. Sci. USA, 86:5673 (1989)). Initially, PCR is carried out with two primers: primer A and primer B, as represented in Figure 2. Primer A contains the RE1 site at its 5' end, followed by poly dG. Primer B is complementary to the constant (C_H1) region of the Cμ gene, 3' to the J region and 5' to primer Z (see Figure 2). Primer B is complementary to all Cμ genes, which encode the heavy chain of molecules of the IgM class, the Ig class expressed by all B cell clones prior to class switching (Schimizu, A. and T. Honjo, Cell, 36:801-803 (1984)) and present in resting B cells. The resultant PCR product includes a significant proportion of cDNAs encompassing the various V_H regions expressed as IgM in the mouse. (The use of other primers complementary to the cDNA genes encoding the constant regions of other immunoglobulin heavy chains can be used in parallel reactions to obtain the variable regions expressed on these molecules, but for simplicity these are not described).

Next, the product of the first PCR procedure is used again for PCR with primer A and primer C. Primer C, like primer B, is complementary to the Cμ gene 3' to J and just 5' to primer B (see Figure 2). Primer C contains the RE2 site at its 5' end. The RE2 sequence is chosen in such a manner that when it is incorporated into the

framework vector, no alteration of coding sequence of the Cμ chain occurs (See Figures 2 and 3). This method of amplifying Cμ cDNAs, referred to as unidirectional nested PCR, incorporates the idea of nested primers for cloning a gene when the nucleotide sequence of only one region of the gene is known (Ohara, O., et al., Proc. Natl. Acad. Sci. USA, 86:5673 (1989)). The PCR product is then cleaved with restriction enzymes RE1 and RE2 and cloned into the RE1 and RE2 sites of the pFHC vector (described below). The sequence of primers and of RE1 and RE2 sites are selected so that when the PCR product is cloned into these sites, the sites are recreated and the cloned antibody gene fragments are brought back into the proper frame with respect to the framework immunoglobulin gene present in pFHC. This results in creation of a Cμ minigene which lacks the intron normally present between J and the C_H1 region of Cμ (See Figure 3). These procedures result in production of the heavy chain library used to produce antigen-binding molecules of the present invention, as described further below.

Optionally, diversity of the heavy chain variable region is increased by random mutagenesis, using techniques known to those of skill in the art.

For example, the library produced as described above is amplified again, using PCR under conditions of limiting nucleotide concentration. Such conditions are known to increase the infidelity of the polymerization and result in production of mutant products. Primers useful for this reaction are Primers C and D, as represented in Figures 2 and 3. Primer D derives from pFHC just 5' to RE1. The PCR product, after cleavage

with RE1 and RE2, is recloned into the framework vector pFHC. To the extent that mutation affects codons of the antigen binding region, this procedure increases the diversity of the binding domains. For example, if the starter library has a complexity of 10^6 elements, and an average of one mutation is introduced per complementarity determining region, and it is assumed that the complementarity determining region is 40 amino acids in size and that any of six amino acid substitutions can occur at a mutated codon, the diversity of the library can be increased by a factor of about 40 x 6, or 240, for single amino acid changes and 240 x 240, or about 6×10^4, for double amino acid changes, yielding a final diversity of approximately 10^{11}. This is considered to be in the range of the diversity of antibodies which animals produce (Tonegawa, S., Nature, 302:575 (1983)). Even greater diversity can be generated by the random combination of H and L chains, the result of co-expression in host cells (see below). It is, thus, theoretically possible to generate a more diverse antibody library in vitro than can be generated in vivo. This library of genes is called the "high diversity" heavy chain library. It may be propagated indefinitely in E. coli. A high diversity light chain library can be prepared similarly.

The framework vector for the light chain library, designated pFLC, includes components similar to those in the vector for the heavy chain library: the enhancer, promoter, a bacterial selectable marker, an animal selectable marker, bacterial origin of replication and light chain exons encoding the constant regions. For

pFLC, the animal selectable marker should differ from the animal selectable in pFHC. For example, if pFHC contains neo^R, pFLC can contain Eco gpt.

A light chain library, which contains diverse light chain fragments, is prepared as described above for construction of the heavy chain library. In constructing the light chain library, the primers used are different from those described above for heavy chain library construction. In this instance, the primers are complementary to light chain mRNA encoding constant regions. The framework vector contains the light chain constant region exons.

Introduction of the Library of Immunoglobulin Chain Genes into Immortalized Animal Cells

The library of immunoglobulin chain genes produced as described is subsequently introduced into a line of immortalized cultured animal cells, referred to as the "host" cells, in which the genes in the library are expressed. Particularly useful for this purpose are plasmacytoma cell lines or myeloma cell lines which have been shown to efficiently express and secrete introduced immunoglobulin genes (Morrison, S.L., et al., Ann. N.Y. Acad. Sci., 507:187 (1987); Kohler, G. and C. Milstein, Eur. J. Immunol., 6:511 (1976); Galfre and C. Milstein, Methods Enzymol., 73:3 (1981); Davis, A.C. and M.J. Shulman, Immunol. Today, 10:119 (1989)). For example, the J558L cell line can be cotransfected using electroporation or protoplast fusion (Morrison, S.L., et al., Ann. N.Y. Acad. Sci., 507:187 (1987)) and transfected

cells selected on the basis of auxotrophic markers present on light and heavy chain libraries.

As a result of cotransformation and selection for markers on both light chain and heavy chain vectors, most transformed host cells will express several copies of immunoglobulin heavy and light chains from the diverse library, and will express chimeric antibodies (antibodies encoded by all or part of two or more genes) (Nisonoff, A., et al., In The Antibody Molecule, Academic Press, NY p. 238 (1975)). These chimeric antibodies are of two types: those in which one chain is encoded by a host cell gene and the other chain is encoded by an exogenously introduced antibody gene and those in which both the light and the heavy chain are encoded by an exogenous antibody gene. Both types of antibodies will be secreted. A library of cells producing antibodies of diverse specificities is produced as a result. The library of cells can be stored and maintained indefinitely by continuous culture and/or by freezing. A virtually unlimited number of cells can be obtained by this process.

Isolation of Cells Producing Antigen-Binding Molecules of Selected Specificity

Cells producing antigen-binding molecules of selected specificity (i.e., which bind to a selected antigen) can be identified and isolated using nitrocellulose filter layering or known techniques. The same methods employed to identify and isolate hybridoma cells producing a desired antibody can be used: cells are pooled and the supernatants tested for reactivity

with antigen (Harlow, E. and D. Lane, Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory, N.Y., p. 283 (1988). Subsequently, individual clones of cells are identified, using known techniques. A preferred method for identification and isolation of cells makes use of nitrocellulose filter overlays, which allow the screening of a large number of cells. Cells from the library of transfected myeloma cells are seeded in 10 cm^2 petri dishes in soft agar (Cook, W.D. and M.D. Scharff, PNAS, 74:5687 (1977); Paige, C.J., et al., Methods in Enzymol., 150:257 (1987)) at a density of 10^4 colony forming units, and allowed to form small colonies (approximately 300 cells). A large number of dishes (>100) may be so seeded. Cells are then overlayed with a thin film of agarose (<1mm) and the agarose is allowed to harden. The agarose contains culture medium without serum. Nitrocellulose filters (or other protein-binding filters) are layered on top of the agarose, and the dishes are incubated overnight. During this time, antibodies secreted by the cells will diffuse through the agarose and adhere to the nitrocellulose filters. The nitrocellulose filters are keyed to the underlying plate and removed for processing.

The method for processing nitrocellulose filters is identical to the methods used for Western blotting (Harlow, E. and D. Lane, Antibodies: Laboratory Manual, Cold Spring Harbor, N.Y., p. 283 (1988)). The antibody molecules are adsorbed to the nitrocellulose filter. The filters, as prepared above, are then blocked. The desired antigen, for example, keyhole lymphet hemocyanin (KLH), which has been iodinated with radioactive ^{125}I, is

then applied in Western blotting buffers to the filters. (Other, non radiographic methods can be used for detection). After incubation, the filters are washed and dried and used to expose autoradiography film according to standard procedures. Where the filters have adsorbed antibody molecules which are capable of binding KLH, the autoradiography film will be exposed. Cells expressing the KLH reactive antibody can be identified by determining the location on the dish corresponding to an exposed filter; cells identified in this manner can be isolated using known techniques. Cells which are isolated from a region of the dish can then be rescreened, to insure the isolation of the clone of antigen-binding molecule-producing cells.

Isolation of Genes Encoding Antigen-Binding Molecules of Selected Specificity and Purification of Encoded Antigen-Binding Molecules

The gene(s) encoding an antigen-binding molecule of selected specificity can be isolated. This can be carried out, for example, as follows: primers D and C (see Figures 2 and 3) are used in a polymerase chain reaction, to produce all the heavy chain variable region genes introduced into the candidate host cell from the library. These genes are cloned again in the framework vector pFHC at the RE1 and RE2 sites. Similarly, all the light chain regions introduced into the host cell from the library are cloned into the light chain vector, pFLC. Members of the family of vectors so obtained are then transformed pairwise into myeloma cells, which are tested for the ability to produce and secrete the antibody with

the desired selectivity. Purification of the antibody from these cells can then be accomplished using standard procedures (Johnstone, A. and R. Thorpe, Immunochem. in Practice, Blackwell Scientific, Oxford, p. 27 (1982); Harlow, E. and D. Lane, Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory, N.Y., p. 283 (1988)).

Alteration of Affinity of Antigen-Binding Molecules

It is also possible to produce antigen-binding molecules whose affinity for a selected antigen is altered (e.g., different from the affinity of a corresponding antigen-binding molecule produced by the present method). This can be carried out, for example, to increase the affinity of an antigen-binding molecule by randomly mutagenizing the genes isolated as described above using previously-described mutagenesis methods. Alternatively, the variable region of antigen-binding molecule-encoding genes can be sequenced and site directed mutagenesis performed to mutate the complementarity determining regions (CDR) (Kabat, E.A., J. Immunol., 141:S 25-36 (1988)). Both processes result in production of a sublibrary of genes which can be screened for antigen-binding molecules of higher affinity or of altered affinity after the genes are expressed in myeloma cells.

Alternative Materials and Procedures for Use in the Present Method

In addition to those described above for use in the method of the present invention, other materials (e.g., starting materials, primers) and procedures can be used

in carrying out the method. For example, use of PCR technology to clone a large collection of cDNA genes encoding variable regions of heavy chains has been described above. Although primers from the Cμ class were described as being used in unidirectional nested PCR, the present invention is not limited to these conditions. For example, primers from any of the other heavy chain classes ($C\gamma_3$, $C\gamma_1$, $C\gamma_{2b}$, Cα for example) or from light chains can be used. Cμ was described as of particular use because of the fact that the entire repertoire of heavy chain variable regions are initially expressed as IgM. Only following heavy-chain class switching are these variable regions expressed with a heavy chain of a different class (Shimizu, A. and T. Honjo, Cell, 36:801-803 (1984)). In addition, the predominant population of B cells in nonimmune spleen cells is IgM^+-cells (Cooper, M.D. and P. Burrows, In Immunoglobulin Genes, Academic Press, N.Y. p. 1 (1989)). Although unidirectional nested PCR amplification is described above, other PCR procedures, as well as other DNA amplification techniques can be used to amplify DNA as needed in the present invention. For example, bidirectional PCR amplification of antibody variable regions can be carried out. This approach requires use of multiple degenerate 5' primers (Orlandi, R., et al., Proc. Natl. Acad. Sci. USA, 86:3833 (1989); Sastry, L., et al., Proc. Natl. Acad. Sci. USA, 86:5728 (1989)). Bidirectional amplification may not pick up the same full diversity of genes as can be expected from unidirectional PCR.

In addition, methods of introducing further diversity into the antibody library other than the method for random mutagenesis utilizing PCR described above can be used. Other methods of random mutagenesis, such as that described by Sambrook, et al. (Sambrook, J., et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989)) can be used, as can direct mutagenesis of the complementarity determining regions (CDRs).

Framework vectors other than one using a mouse $C\mu$ heavy chain constant region, which contains the $C\mu$ enhancer and introns and a viral promoter (described previously) can be used for inserting the products of PCR. The vectors described were chosen for their subsequent use in the expression of the antibody genes, but any eukaryotic or prokaryotic cloning vector could be used to create a library of diverse cDNA genes encoding variable regions of antibody molecules. The inserts from this vector could be transferred to any number of expression vectors. For example, other framework vectors which include intronless genes can be constructed, as can other heavy chain constant regions. In addition to plasmid vectors, viral vectors or retroviral vectors can be used to introduce genes into myeloma cells.

The source for antibody molecule mRNAs can also be varied. Purified resting B lymphocytes from mouse nonimmunized spleen are described above as such a source. However, total spleens (immunized or not) from other animals, including humans, can be used, as can any source of antibody-producing cells (e.g., peripheral blood, lymph nodes, inflammatory tissue, bone marrow).

Introduction of H and L chain gene DNA into myeloma cells using cotransformation by electroporation or protoplast fusion methods is described above (Morrison, S.L. and V.T. Oi, Adv. Immunol., 44:65 (1989)). However, any means by which DNA can be introduced into living cells in vivo can be used, provided that it does not significantly interfere with the ability of the transformed cells to express the introduced DNA. In fact, a method other than cotransformation, can be used. Cotransfection was chosen for its simplicity, and because both the H and L chains can be introduced into myeloma cells. It may be possible to introduce only the H chain into myeloma cells. Moreover, the H chain itself in many cases carries sufficient binding affinity for antigen. However, other methods can also be used. For example, retroviral infection may be used. Replication-incompetent retroviral vectors can be readily constructed which can be packaged into infective particles by helper cells (Mann, R., et al., Cell, 33:153-159 (1903)). Viral titers of 10^5 infectious units per ml. can be achieved, making possible the transfer of very large numbers of genes, into myeloma cells.

Further increases in the diversity of antibody-producing cells than results from the method described above can be generated if light and heavy chain genes are introduced separately into myeloma cells. Light chain genes can be introduced into one set of myeloma cells with one selectable marker, and heavy chains into another set of cells with a different selectable marker. Myeloma cells containing and expressing both H and L chains could then be generated by the highly efficient process of

polyethylene glycol mediated cell fusion (Pontecorvo, G., Somatic Cell Genetics, 1:397 (1975)). Thus, a method of screening diverse libraries of antibody genes using animal cells is not limited by the number of cells which can be generated, but by the number of cells which can be screened.

Methods of identifying antigen-binding molecule-expressing cells expressing an antigen-binding molecule of selected specificity other than the nitrocellulose filter overlay technique described above can be used. An important characteristic of any method is that it be useful to screen large numbers of different antibodies. With the nitrocellulose filter overlay technique, for example, if 300 dishes are prepared and 10^4 independent transformed host cells per dish are screened, and if, on average, each cell produces ten different antibody molecules, then 300 x 10^4 x 3, or about 10^7 different antibodies can be screened at once. However, if the antibody molecules can be displayed on the cell surface, still larger numbers of cells can be screened using affinity matrices to pre-enrich for antigen-binding cells. There are immortal B cell lines, such as BCL_1B_1, which will express IgM both on the cell surface and as a secreted form (Granowicz, E.S., et al., J. Immunol., 125:976 (1980)). If such cells are infected by retroviral vectors containing the terminal Cμ exons, the infected cells will likely produce both secreted and membrane bond forms of IgM (Webb, C.F., et al., J. Immunol., 143:3934-3939 (1989)). Still other methods can be used to detect antibody production. If the host cell is E. coli, a nitrocellulose overlay is possible, and

such methods have been frequently used to detect E. coli producing particular proteins (Sambrook, J., et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989)). Other methods of detection are possible and one in particular, which involves the concept of "viral coating", is discussed below.

Viral coating can be used as a means of identifying viruses encoding antigen-combining molecules. In this method, a viral vector is used to direct the synthesis of diverse antibody molecules. Upon lytic infection of host cells, and subsequent cell lysis, the virus becomes "coated" with the antibody product it directs. That is, the antibody molecule becomes physically linked to the outside of a mature virus particle, which can direct its synthesis. Methods for viral coating are described below. Viruses coated by antibody can be physically selected on the basis of their affinity to antigen which is attached to a solid support. The number of particles which can be screened using this approach is well in excess of 10^9 and it is possible that 10^{11} different antibody genes could be screened in this manner. In one embodiment, an affinity matrix containing antigen used to purify those viruses encoding antibody molecules with affinity to antigen and which coat the surface of the virus which encodes those antibodies is used.

One method of viral coating is as follows: A diverse library of bacteriophage λ encoding parts of antibody molecules that are expressed in infected E. coli and which retain the ability to bind antigens is created,

using known techniques (Orlandi, R., et al., Proc. Natl. Acad. Sci. USA, 86:3833 (1989); Huse, W.D., et al., Science, 246:1275 (1989); Better, M., et al., Science, 240:1041 (1988); Skerra, A. and A. Pluckthon, Science, 240:1038 (1988)). Bacteria infected with phage are embedded in a thin film of semisolid agar. Greater than 10^7 infected bacteria may be plated in the presence of an excess of uninfected bacteria in a volume of 1 ml of agar and spread over a 10 cm^2 surface. The agar contains monovalent antibody "A" (Parham, P., In Handbook of Experimental Immunology: Immunochem., Blackwell Scientific Publishers, Cambridge, MA, pp. 14.1-14.23 (1986)), which can bind the λ coat proteins and which has been chemically coupled to monovalent antibody "B", which can bind an epitope on all viral directed antibody molecules. Monovalent antibodies are used to prevent the crosslinking of viral particles. Upon lytic burst, progeny phage particles become effectively cross linked to the antibody molecule they encode. Because lysis occurs in semisolid medium, in which diffusion is slow, cross linking between a given phage and the antibody encoded by another phage is minimized. A nitrocellulose filter (or other protein binding filter) is prepared as an affinity matrix by adsorbing the desired antigen. The filter is then blocked so that no other proteins bind nonspecifically. The filter is overlayed upon the agar, and coated phage are allowed to bind to the antigen by way of their adherent antibody molecules. Filters are washed to remove nonspecifically bound phage. Specifically bound phage therefore represent phage

encoding antibodies with the desired specificity. These can now be propagated by reinfection of bacteria.

Thus the present invention makes it possible to produce antigen-binding molecules which, like antibodies produced by presently-available techniques, bind to a selected antigen (i.e., having binding specifity). Antibodies produced as described can be used, for example, to detect and neutralize antigens and deliver molecules to antigenic sites.

EXAMPLE I Amplification of IgM Heavy Chain Variable Region DNA from mRNA

IgM heavy chain variable DNA is amplified from mRNA by the procedure represented schematically in Figure 2. In Figure 2, Panel A depicts the relevant regions of the poly adenylated mRNA encoding the secreted form of the IgM heavy chain. In Panel A, S denotes the sequences encoding the signal peptide which causes the nascent peptide to cross the plasma membrane, a necessary step in the processing and secretion of the antibody. V, D and J derive from separate exons and together comprise the variable region. C_H1, C_H2, and C_H3 are the three constant domains of $C\mu$. "Hinge" encodes the hinge region. C, B and Z are oligonucleotide PCR primers used in the amplification process. The only constraints on Primers B and Z are that they are complementary to the mRNA, and occur in the order shown relative to C. Primer C, in addition to being complementary to mRNA, has an extra bit of sequence at its 5' end which allows the cloning of its PCR product. This is described below. Panel B depicts the reverse transcript DNA product of the mRNA primed by

oligonucleotide Z, with the addition of poly-dC by terminal transferase at the 3' end of the product. Panel C depicts the annealing of primer A to the reverse transcript DNA represented in Panel B. Primer A contains the restriction endonuclease site RE1, with additional DNA at its 5' end. The constraints on the RE1 site are described in Example 2. Panel D depicts the final double stranded DNA PCR product made utilizing primers A and B. Panel E depicts the PCR product shown in Panel D annealed to Primer C. Panel F is a blow up of panel E showing the structure of primer C. Primer C consists of two parts: a 3' part complementary to IgM heavy chain mRNA as shown, and a 5' part which contains restriction site RE2 and spacer. Constraints on RE2 are described in Example 2. Panel G depicts the final double stranded DNA PCR product utilizing Primers A and C and the product of the previous PCR (depicted in Panel D) as template. The S, V, D, J regions are again depicted.

EXAMPLE 2 Construction of Heavy Chain Framework Vector pFHC

A heavy chain framework vector, designated pFHC, is constructed, using known techniques (See Figure 3). It is useful for introducing antibody-encoding DNA into host cells, in which the DNA is expressed, resulting in antibody production. The circular plasmid (above) is depicted linearized (below) and its relevant components are shown. The neomycin antibiotic resistance gene (neo^R) is useful for selecting transformed animal cells (Sambrook, J., et al., Molecular Cloning: A Laboratory Manual, 2d Ed., Cold Spring Harbor Laboratory Press, Cold

Spring Harbor, NY (1989)). The bacterial replication origin and ampicillin antibiotic resistance genes, useful respectively, for replication in E. coli and rendering E. coli resistant to ampicillin, can derive from any number of bacterial plasmids, including PBR322 (Sambrook, J., et al., Molecular Cloning: A Laboratory Manual, 2d Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1989)). The Cμ enhancer, which derives from the intron between exons J and C_H1 of the Cμ gene, derives from any one of the cloned Cμ genes (Kawakami, T., et al., Nucleic Acids Research, 8:3933 (1980); Honjo, T., Ann. Rev. Immunol., 1:499 (1983)) and increases levels of transcription from antibody genes. LTR contains the viral promoter from the Moloney MLV retrovirus DNA (Mulligan, R.C., Experimental Manipulation of Gene Expression, New York Academic Press, p. 155 (1983)). D represents the PCR primer described in the text, depicted in its 5' to 3' orientation. The only constraints on D are its orientation, its complementarity to pFHC and its order relative to the RE1 and RE2 cloning sites. Preferably, D is within 100 nucleotides of RE1. The cDNA cloning site contains restriction endonuclease sites RE1 and RE2, separated by spacer DNA which allows their efficient cleavage. The constraints on RE1 and RE2 are described below. The Cμ exons, as described in the text and literature, direct the synthesis of IgM heavy chain. Only part of C_H1 is present, as described below. C_H3 is chosen to contain the Cμs region which specifies a secreted form of the heavy chain ((Kawakami, T., et al., Nucleic Acids Research, 8:3933 (1980); Honjo, T., Ann. Rev. Immunol., 1:499 (1983)). Finally, pFHC contains

poly A addition and termination sequences which can be derived from the Cμ gene itself (Honjo, T., Ann. Rev. Immunol., 1:499 (1983); Kawakami, T., et al., Nucleic Acids Research, 8:3933 (1980)). One potential advantage of using the entire Cμ gene is that in some host cell systems, a membrane bound and secreted form of IgM may be expressed (Granowicz, E.S., et al., J. Immunol, 125:976 (1980)).

The plasmid can be produced by combining the individual components, or nucleic acid segments, depicted in Figure 3, using PCR cassett assembly (See below). Because the entire nucleotide sequence of each component is defined, the entire nucleotide sequence of the plasma is defined.

The constraints on RE1 are simple. It should be the sole cleavage site on the plasmid for its restriction endonuclease. The choice of RE1 can be made by computer based sequence analysis (Intelligenetics Suite, Release 5:35, Intelligenetics).

The constraints on RE2 are more complex. First, it must be the sole cleavage site on the plasmid for its restriction endonuclease, as described for RE1. Moreover, the RE2 site must be such that when the PCR product is inserted, a gene is thereby created which is capable of directing the synthesis of a complete IgM heavy chain. This limits the choices for RE2, but the choices available can be determined by computer based sequence analysis. The choices can be determined as follows. First, a list of restriction endonucleases that do not cleave pFHC is compiled (see Table 1).

TABLE 1

Non-Cutting Enzymes for the Mouse Cμ Gene

AatII	AhaII	AseI
AvrII	BgII	BspHI
BssHII	BstBI	ClaI
DraI	EagI	EcoRI
EcoRV	FspI	HgaI
HincII	HpaI	KpnI
MluI	NaeI	NarI
NdeI	NotI	NruI
PaeR7I	PvuI	RsrII
SacII	SaII	ScaI
SfII	SnaBI	SpeI
SphI	SspI	StuI
Tth111I	XbaI	XhoI

These are called the "rare non-cutters." Next, the sequence of C_H1 is rewritten with "N" at the third position of each codon and entered into the computer. This is called the "N-doped sequence" (See Figure 4). Next, the rare non-cutters are surveyed by computer analysis for those which will cleave the N-doped sequence. The search program will show a possible restriction endonuclease site, assuming a match between N and the restriction endonuclease cutting site. For example, with 39 rare non-cutters, 22 will cleave the N-doped sequence of Cμ C_H1, many of them several times (see Table 2). In this table, "Def" means a definite cut

site, of which there are none, because of the Ns. "Pos" means a possible cleavage site at the indicated nucleotide position if N is chosen appropriately. "Y" indicates any pyrimidine, "R" indicates any purine and "N" indicates any nucleotide. The nucleotide positions refer to coordinates represented in Figure 4.

TABLE 2

ENZYME	RECOGNITION	CUT SITE
AatII	(GACGTC)	Def : none Pos : 250 309
AhaII	(GRCGYC)	Def : none Pos : 247 306
AvrII	(CCTAGG)	Def : none Pos : 204
BspHI	(TCATGA)	Def : none Pos : 138
BsshII	(GCGCGC)	Def : none Pos : 189
EcoRI	(GAATTC)	Def : none Pos : 195 334
EcoRV	(GATATC)	Def : none Pos : 214
HgaI	(GACGCNNNNN) (NNNNNNNNNNGCGTC)	Def : none Pos : 284
HincII	(GTYRAC)	Def : none Pos : 183 220
HpaI	(GTTAAC)	Def : none Pos : 220
KpnI	(GGTACC)	Def : none Pos : 408
NruI	(TCGCGA)	Def : none Pos : 174 193 303
PaeR7	(CTCGAG)	Def : none Pos : 190 339
PvuI	(CGATCG)	Def : none Pos : 178
ScaI	(AGTACT)	Def : none Pos : 209 266 284
SpeI	(ACTAGT)	Def : none Pos : 131 167 359
SphI	(GCATGC)	Def : none Pos : 338
SspI	(AATATT)	Def : none Pos : 371
StuI	(AGGCCT)	Def : none Pos : 149
Tth111I	(GACNNNGTC)	Def : none Pos : 212
XbaI	(TCTAGA)	Def : none Pos : 338
XhoI	(CTCGAG)	Def : none Pos : 190 339

Most of these cleavage sites (about 60%) are compatible with the amino acids specified by C_H1. Therefore, it is possible to mutate C_H1 to create a unique site for such an enzyme without altering the amino acid sequence incoded by C_H1. One sequence which illustrates this is shown below:

1)	...ala	met	gly	cys	leu	ala	arg	asp...
2)	...GCC	ATG	GGC	TGC	CTA	GCC	CGG	GAC...
3)	...GCC	ATG	GGC	TGC	CTA	GCG	CGC	GAC...
						BssHII		

Line 1 represents part of the actual amino acid sequence specified by the mouse $C\mu$ C_H1 gene region, and line 2 is the actual nucleotide sequence. By changing the sequence to the indicated nucleotides underlined on line 3, a cleavage site for the rare non-cutter BssHII is created. The new sequence (containing the BssHII site) GCG CGC still encodes the identical amino acid sequence. Therefore, the sequence of the primer C is chosen to be the complement of line 3, and RE2 is the BssHII site. Such a primer will function in the PCR and vector construction as desired. Other examples are possible, and the same process can be used in designing vectors and primers for cloning light chain variable regions.

The choice for primer C puts a constraint on pFHC. In the example shown, the C_H1 region contained on pFHC must begin at its 5' end with the mutant sequence GCG CGC. Such mutant fragments can be readily made by the process of PCR cassette assembly described below.

The process of PCR cassette assembly is a method of constructing plasmid molecules (in this case the plasmid pFHC) from fragments of DNA of known nucleotide sequence. One first compiles a list of restriction endonucleases that do not cleave any of the fragments. Each fragment is then individually PCR amplified using synthesized oligonucleotide primers complementary to the terminal sequences of the fragment. These primers are synthesized to contain on their 5' ends restriction endonuclease cleavage sites from the compiled list. Thus, each PCR product can be so designed that each fragment can be assembled one by one into a larger plasmid structure by cleavage and ligation and transformation into E. coli. Using this method, it is also possible to make minor modifications to modify the terminal sequence of the fragment being amplified. This is done by altering the PCR primer slightly so that a mismatch occurs. In this way it is possible to amplify the $C\mu$ gene starting precisely from the desired point in C_H1 (as determined by oligo C above) and creating the RE2 endonuclease cleavage site.

CLAIMS

1. An in vitro process for synthesizing DNA encoding a family of antigen-combining proteins, comprising the steps of:
   a) obtaining DNA containing genes encoding antigen-combining proteins;
   b) combining the DNA containing genes encoding antigen-combining proteins with sequence specific primers which are oligonucleotides homologous to conserved regions of the genes; and
   c) performing sequence specific gene amplification.

2. DNA encoding a family of antigen-combining proteins produced by the process of Claim 1.

3. The process of Claim 1 wherein sequence specific gene amplification is performed by the polymerase chain reaction.

4. The process of Claim 3 wherein the sequence specific primers are bidirectional.

5. The process of Claim 3 wherein the sequence specific primers are nested unidirectional primers.

6. The process of Claim 1 wherein the antigen-combining proteins are immunoglobulins.

7. The process of Claim 6 wherein the immunoglobulins are selected from the group consisting of heavy chains and light chains.

8. The process of Claim 7 wherein the heavy chains are μ chains.

9. The process of Claim 1 wherein the DNA containing genes encoding antigen-combining proteins is cDNA of RNA from antibody-producing cells.

10. The process of Claim 1 wherein the DNA containing genes encoding antigen-combining proteins is genomic DNA from antibody-producing cells.

11. The process of Claim 8 wherein the antigen-combining proteins are of mammalian origin.

12. The process of Claim 1 wherein the primers are oligonucleotides homologous to conserved regions of the constant regions of immunoglobulin genes.

13. The process of Claim 1 wherein the primers are oligonucleotides homologous to the conserved regions of the variable regions of immunoglobulin genes.

14. The process of Claim 1 wherein the primers contain at least one restriction endonuclease cloning site.

15. The process of Claim 1 wherein the primers are selected from the group consisting of oligonucleotide B of Figure 2 and oligonucleotide C of Figure 2.

16. A method of creating a diverse starter library of DNAs encoding families of antigen-combining proteins comprising cloning the product of Claim 1 into an appropriate vector.

17. A diverse starter library of DNAs encoding families of antigen-combining proteins produced by the method of Claim 14.

18. The method of Claim 16 wherein the vector is a prokaryotic vector or a eukaryotic vector.

19. The method of Claim 16 wherein the vector is a viral vector or a retroviral vector.

20. The method of Claim 16 wherein the vector is a plasmid.

21. The method of Claim 20 wherein the plasmid is selected from the group consisting of pFHC and pLHC.

22. The method of Claim 16 wherein the vector is selected from the group consisting of expression vectors and cloning vectors.

23. The method of Claim 22 wherein the expression vector is appropriate for expression of the variable region of an antigen-combining protein as a chimeric molecule in register with a framework protein.

24. The method of Claim 23 wherein the framework protein is an immunoglobulin.

25. The method of Claim 24 wherein the immunoglobulin is all or a portion of the constant region of the μ heavy chain.

26. The method of Claim 16 further comprising creating a collection of viral particles from viral vector-based libraries of DNA encoding antigen-combining proteins by the process of introducing viral vectors into host cells in which they replicate and form viral particles.

27. A method of producing a high diversity library of DNA encoding families of antigen-combining proteins comprising mutagenizing the product of Claim 16.

28. A high diversity library of DNA encoding families of antigen-combining proteins produced by the method of Claim 27.

29. The method of Claim 27 wherein mutagenizing is carried out by random chemical mutagenesis.

30. The method of Claim 27 wherein mutagenizing is carried out by performing the polymerase chain reaction under limiting nucleotide conditions.

31. The method of Claim 27 wherein mutagenizing is carried out in such a manner that mutagenesis is limited to DNA encoding variable regions of the antigen-combining protein.

32. A process of producing a diverse population of host cells which comprises introducing into host cells

DNA of the starter library or high diversity libraries of antigen-combining proteins.

33. Host cells produced by the method of Claim 32.

34. The process of Claim 32 wherein the host cells are prokaryotic.

35. The process of Claim 32 wherein the host cells are eukaryotic.

36. The process of Claim 35 wherein the host cells are selected from the group consisting of immortalized cultured mammalian cells.

37. The process of Claim 36 wherein the immortalized cultured mammalian cells are selected from the group consisting of myelomas and plasmacytomas.

38. The process of Claim 32 wherein the libraries encoding families of antigen-combining proteins are introduced into host cells by a method selected from the group consisting of: electroporation, calcium phosphate coprecipitation, protoplast fusion, viral infection, and cell fusion.

39. The process of Claim 32 wherein the libraries of DNAs encoding families of antigen-combining proteins is contained in an expression vector.

40. The process of Claim 32 wherein the DNAs encoding families of antigen-combining proteins encode antigen-combining proteins selected from the group

consisting of immunoglobulin heavy chain variable regions or immunoglobulin light chain variable regions.

41. The process of Claim 40 wherein DNAs encoding immunoglobulin heavy chain variable regions are introduced simultaneously with or sequentially to DNAs encoding immunoglobulin light chain variable regions.

42. The method of Claim 32 further comprising identifying cells which produce antigen-combining molecules of selected specificity.

43. The method of Claim 42 wherein identifying of cells which produce antigen-combining molecules of selected specificity is carried out by assaying cellular supernatants for antigen-combining activity.

44. The method of Claim 42 wherein identifying of cells which produce antigen-combining molecules of selected specificity is carried out by a nitrocellulose filter overlay technique.

45. The method of Claim 44 wherein cells producing antigen-combining molecules of selected specificity are enriched for cells producing antigen-combining molecules on their surface by affinity matrix chromatography.

46. Cells produced by the method of Claim 42.

47. Antigen-combining molecules produced by cells of Claim 42.

48. DNAs encoding immunoglobulin heavy chain variable regions or immunoglobulin light chain variable regions, present in cells of Claim 42.

49. Viruses produced by the method of Claim 26.

50. A method of isolating viruses of Claim 49 encoding antigen-combining molecules of selected specificity, comprising the steps of:

a) infecting host cells with an appropriate virus containing DNA encoding antigen-combining molecules;

b) coating the virus with antigen-combining molecules which the virus encodes; and

c) subjecting the product of step (b) to affinity-matrix selection, to separate the virus according to the antigen-combining molecules they contain.

51. Viruses produced by the method of Claim 50.

52. Antigen-combining molecules encoded by viruses of Claim 51.

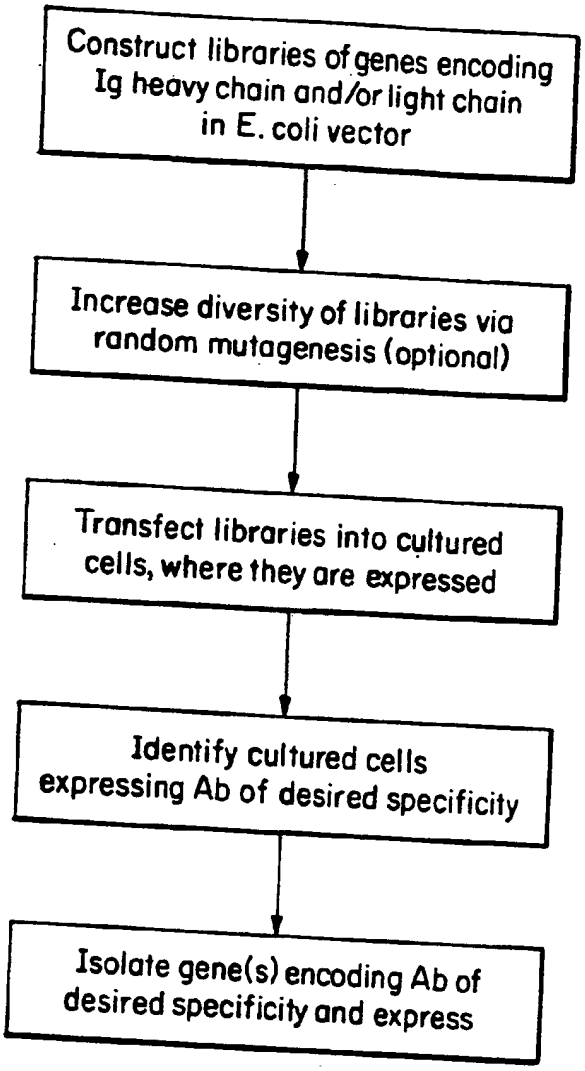
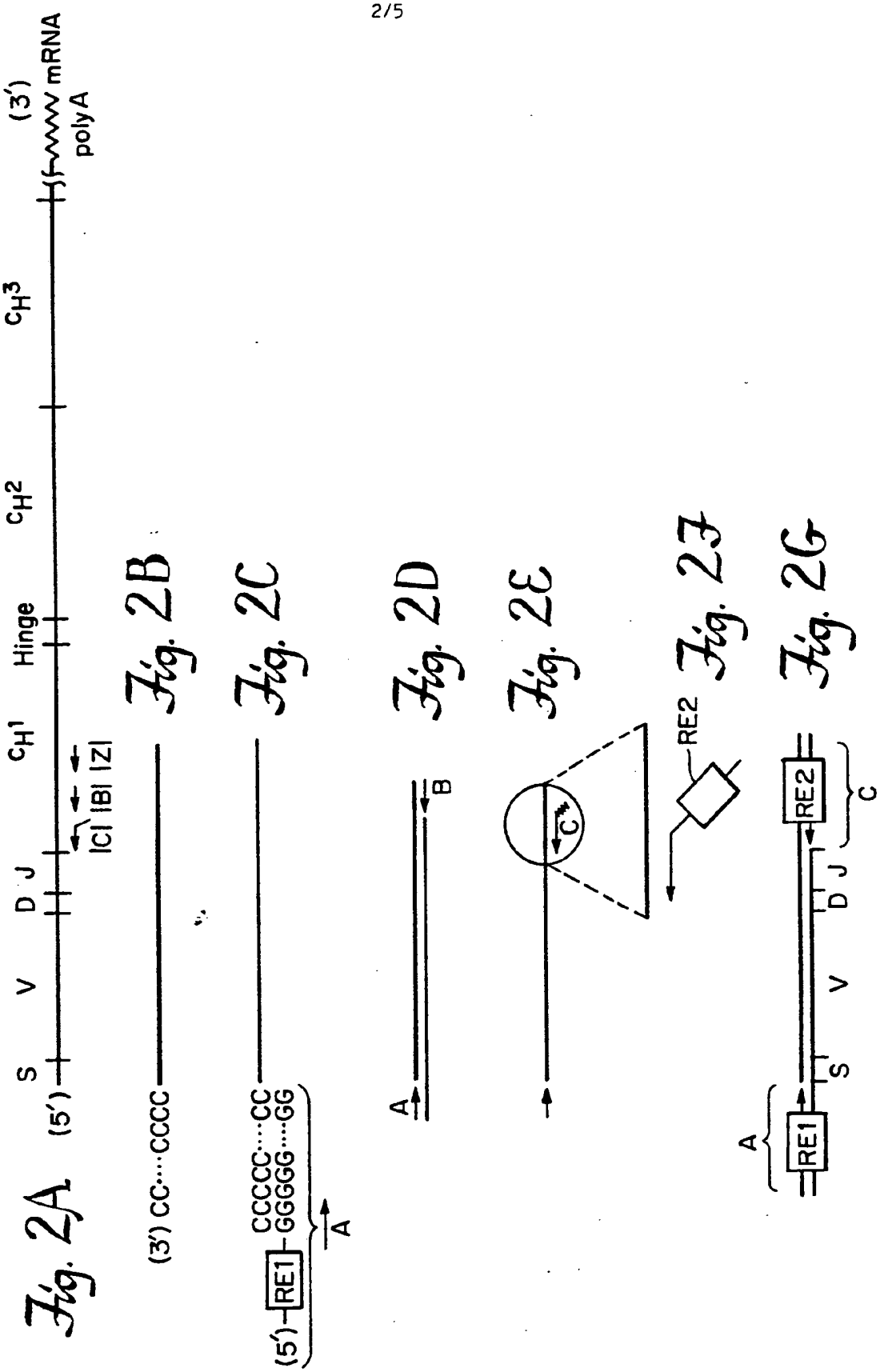

Construct libraries of genes encoding Ig heavy chain and/or light chain in E. coli vector
Increase diversity of libraries via random mutagenesis (optional)
Transfect libraries into cultured cells, where they are expressed
Identify cultured cells expressing Ab of desired specificity
Isolate gene(s) encoding Ab of desired specificity and express

Fig. 1


(5')
S
V
D J
|C| |B| |Z|
C_H1
Hinge
C_H2
C_H3
(3')
mRNA
polyA
(3') CC····CCCC
CCCCC····CC
(5') RE1 GGGGG····GG
A
B
C
RE2
RE1
RE2
A
C
Fig. 2A
Fig. 2B
Fig. 2C
Fig. 2D
Fig. 2E
Fig. 2F
Fig. 2G

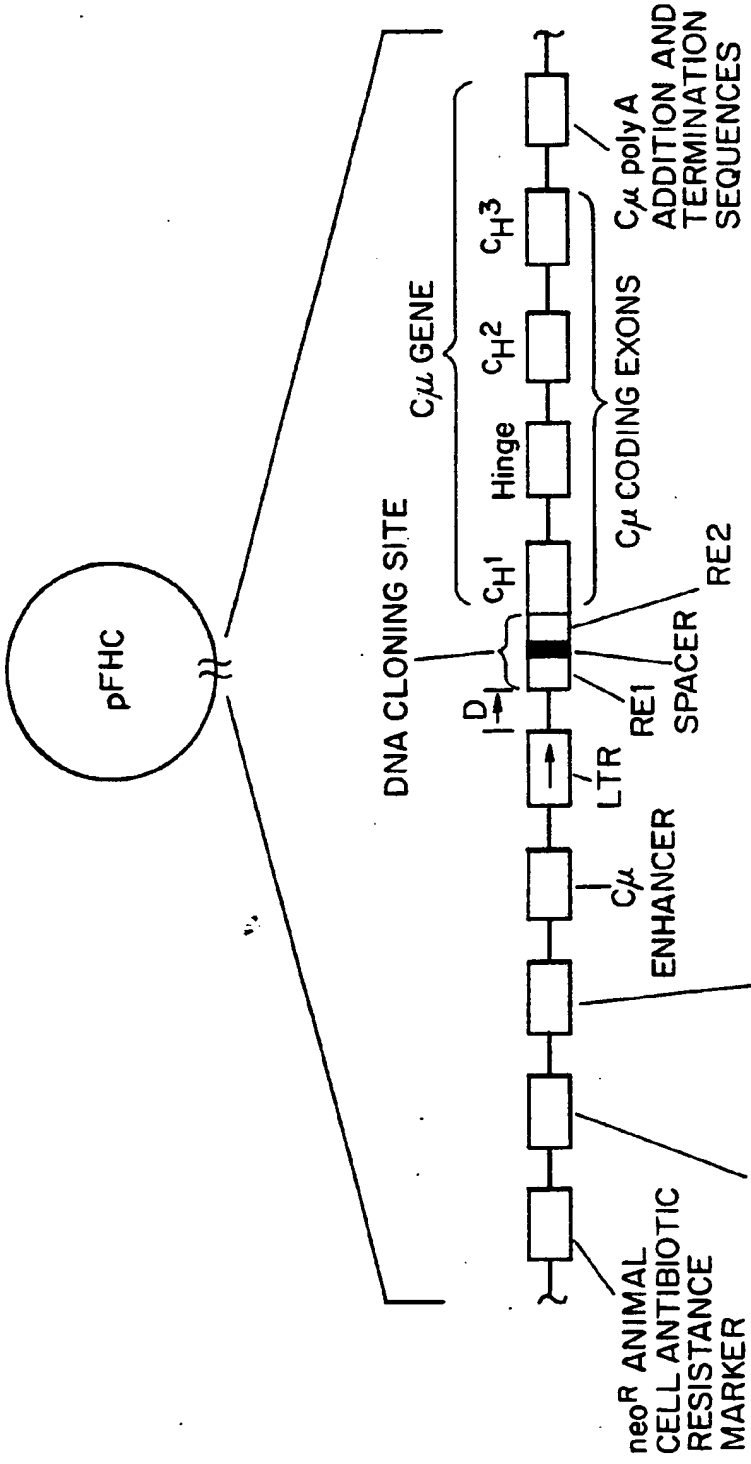

pFHC
DNA CLONING SITE
Cμ GENE
C_H1
Hinge
C_H2
C_H3
Cμ CODING EXONS
Cμ poly A ADDITION AND TERMINATION SEQUENCES
RE2
SPACER
RE1
D
LTR
Cμ ENHANCER
AMPICILLIN RESISTANCE BACTERIAL SELECTION MARKER
BACTERIAL REPLICATION ORIGIN
neo^R ANIMAL CELL ANTIBIOTIC RESISTANCE MARKER

Fig. 3

									>		109				118		127
									AGT	CAG	TCC	TTC	CCA	AAT	GTC	TTC	
									Ser	Gln	Ser	Phe	Pro	Asn	Val	Phe	
		136			145			154			163			172			181
CCC	CTC	GTC	TCC	TGC	GAG	AGC	CCC	CTG	TCT	GAT	AAG	AAT	CTG	GTG	GCC	ATG	GGC
Pro	Leu	Val	Ser	Cys	Glu	Ser	Pro	Leu	Ser	Asp	Lys	Asn	Leu	Val	Ala	MET	Gly
		190			199			208			217			226			235
TGC	CTA	GCC	CGG	GAC	TTC	CTG	CCC	AGC	ACC	ATT	TCC	TTC	ACC	TGG	AAC	TAC	CAG
Cys	Leu	Ala	Arg	Asp	Phe	Leu	Pro	Ser	Thr	Ile	Ser	Phe	Thr	Trp	Asn	Tyr	Gln
		244			253			262			271			280			289
AAC	AAC	ACT	GAA	GTC	ATC	CAG	GGT	ATC	AGA	ACC	TTC	CCA	ACA	CTG	AGG	ACA	GGG
Asn	Asn	Thr	Glu	Val	Ile	Gln	Gly	Ile	Arg	Thr	Phe	Pro	Thr	Leu	Arg	Thr	Gly
		298			307			316			325			334			343
GGC	AAG	TAC	CTA	GCC	ACC	TCG	CAG	GTG	TTG	CTG	TCT	CCC	AAG	AGC	ATC	CTT	GAA
Gly	Lys	Tyr	Leu	Ala	Thr	Ser	Gln	Val	Leu	Leu	Ser	Pro	Lys	Ser	Ile	Leu	Glu
		352			361			370			379			388			397
GGT	TCA	GAT	GAA	TAC	CTG	GTA	TGC	AAA	ATC	CAC	TAC	GGA	GGC	AAA	AAC	AGA	GAT
Gly	Ser	Asp	Glu	Tyr	Leu	Val	Cys	Lys	Ile	His	Tyr	Gly	Gly	Lys	Asn	Arg	Asp
		406			415	>											
CTG	CAT	GTG	CCC	ATT	CCA	G											
Leu	His	Val	Pro	Ile	Pro												

Fig. 4A

```
                                                             109            118            127
                                                         >   AGN  CAN  TCN  TTN  CCN  AAN  GTN  TTN
                                                                       Ser       Pro       Val

               136            145            154            163            172            181
CCN  CTN  GTN  TCN  TGN  GAN  AGN  CCN  CTN  TCN  GAN  AAN  AAN  CTN  GTN  GCN  ATN  GCN
Pro  Leu  Val  Ser                 Pro  Leu  Ser                 Leu  Val  Ala       Gly

               190            199            208            217            226            235
TGN  CTN  GCN  CGN  GAN  TTN  CTN  CCN  AGN  ACN  ATN  TCN  TTN  ACN  TGN  AAN  TAN  CAN
     Leu  Ala  Arg            Leu  Pro       Thr       Ser       Thr

               244            253            262            271            280            289
AAN  AAN  ACN  GAN  GTN  ATN  CAN  GGN  ATN  AGN  ACN  TTN  CCN  ACN  CTN  AGN  ACN  GGN
          Thr       Val            Gly            Thr       Pro  Thr  Leu       Thr  Gly

               298            307            316            325            334            343
GGN  AAN  TAN  CTN  GCN  ACN  TCN  CAN  GTN  TTN  CTN  TCN  CCN  AAN  AGN  ATN  CTN  GAN
Gly            Leu  Ala  Thr  Ser       Val       Leu  Ser  Pro                 Leu

               352            361            370            379            388            397
GGN  TCN  GAN  GAN  TAN  CTN  GTN  TGN  AAN  ATN  CAN  TAN  GGN  GGN  AAN  AAN  AGN  GAN
Gly  Ser                 Leu  Val                           Gly  Gly

               406            415
CTN  CAN  GTN  CCN  ATN  CCN  > G
Leu       Val  Pro       Pro
```

Fig. 4B

INTERNATIONAL SEARCH REPORT

International Application No PCT/US 91/00209

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) [6]

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC[5]: C 12 N 15/13, C 07 K 15/28

II. FIELDS SEARCHED

Minimum Documentation Searched [7]

Classification System	Classification Symbols
IPC[5]	C 07 K, C 12 N

Documentation Searched other than Minimum Documentation to the Extent that such Documents are included in the Fields Searched [8]

III. DOCUMENTS CONSIDERED TO BE RELEVANT [9]

Category [9]	Citation of Document, [11] with indication, where appropriate, of the relevant passages [12]	Relevant to Claim No. [13]
X	Science, vol. 246, 8 December 1989, W.D. Huse et al.: "Generation of a large combinatorial library of the immunoglobulin repertoire in phage lambda", pages 1275-1281 see the whole article; especially tables 1 and 2 --	1-3,5-14, 16-20,21-25
X	Proceedings of the National Academy of Sciences, vol. 86, August 1989, L. Sastry et al.: "Cloning of the immunological repertoire in Escherichia coli for generation of monoclonal catalytic antibodies: construction of a heavy chain variable region-specific cDNA library", pages 5728-5732 see the whole article --	1-3,5-14, 16-20,21-25

./.

* Special categories of cited documents: [10]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search	Date of Mailing of this International Search Report
16th May 1991	25/06/91
International Searching Authority	**Signature of Authorized Officer**
EUROPEAN PATENT OFFICE	Danielle van der Haas

Form PCT/ISA/210 (second sheet) (January 1985)

International Application No PCT/US 91/00209

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)

Category *	Citation of Document, [11] with indication, where appropriate, of the relevant passages	Relevant to Claim No.
X	Nature, vol. 341, no. 6242, 12 October 1989, (London, GB), E.S. Ward et al.: "Binding activities of a repertoire of single immunoglobulin variable domains secreted from Escherichia coli", pages 544-546 see the whole article --	1-3,5-14, 16-20,21-25
X	Proceedings of the National Academy of the USA, vol. 86, no. 10, May 1989, (Washington, DC, US), R. Orlandi et al.: "Cloning immunoglobulin variable domains for expression by the polymerase chain reaction", pages 3833-3837 see the whole article --	1-3,5-14, 16-20,21-25
A	Clinical Chemistry, vol. 35, no. 9, September 1989, G.P. Moore: "Genetically engineered antibodies", pages 1849-1853 see page 1852, left-hand column	1-52
